

ENTWICKLUNG, VALIDIERUNG UND
EINSATZ EINER SIMULIERTEN
KLASSENRAUMUMGEBUNG ZUR
ERFASSUNG
EVOLUTIONSBEZOGENEN
PROFESSIONSWISSENS

DISSERTATION

ZUR ERLANGUNG DES DOKTORGRADES DER
MATHEMATISCH-NATURWISSENSCHAFTLICHEN FAKULTÄT DER
CHRISTIAN-ALBRECHTS-UNIVERSITÄT ZU KIEL

Dr. rer. nat.

VORGELEGT VON

JULIAN FISCHER

KIEL, FEBRUAR 2021

Erste Gutachterin

Prof. Dr. Ute Harms

Zweiter Gutachter

Prof. Dr. Jens Möller

Tag der mündlichen Prüfung

7. Mai 2021

Die Studien im Rahmen dieser Dissertation wurden vom Bundesministerium für Bildung und Forschung (BMBF) gefördert [16DHL1014].

ZUSAMMENFASSUNG

Das Professionswissen gilt als Kern professioneller Kompetenz und ist die zentrale Voraussetzung für einen qualitätvollen Unterricht sowie den damit verbundenen Schülerleistungen. Grundlegend kann das Professionswissen in Wissensbereiche – fachliches, fachdidaktisches und pädagogisches Wissen – sowie in Wissenstypen – deklaratives und prozedurales Wissen – differenziert werden. In der bisherigen Forschung wurden zahlreiche Studien durchgeführt, um das Professionswissen von Lehrkräften systematisch zu untersuchen. Hierbei mangelte es an Messinstrumenten, welche nicht nur deklaratives, sondern auch prozedurales Professionswissen erfassen können. Eine Lösung stellen simulierte Klassenraumumgebungen dar, in welchen die angehenden Lehrkräfte das im akademischen Diskurs erlernte deklarative Wissen handlungsnah anwenden können. Im Biologieunterricht wird prozedurales Wissen vor allem dann relevant, wenn es um die Beurteilung von Schülerleistungen geht. Besonders im Kontext der biologischen Evolution, welche den integrativen Rahmen des Biologieunterrichts darstellt, werden häufig Fehlvorstellungen zur Erklärung evolutionsbiologischer Prozesse geäußert.

Die vorliegende Dissertation beinhaltet drei empirische Studien. Alle Studien untersuchen Facetten des prozeduralen Professionswissens mit einem Fokus auf dem biologiedidaktischen Wissen. In Studie 1 wird ein digitales Instrument – das Schülerinventar – eingesetzt, um zu untersuchen, inwiefern virtuelle Schülerklausuren auf Basis des Professionswissens von Lehrerinnen und Lehrern im Vorbereitungsdienst beurteilt werden können. Hier zeigten sich Defizite im prozeduralen fachdidaktischen sowie pädagogischen Wissen. In Studie 2 wird eine im Rahmen dieser Dissertation weiterentwickelte simulierte Klassenraumumgebung – der Simulierte Klassenraum Biologie (SKR^{Bio}) – erprobt und validiert. Die Ergebnisse zeigten, dass der SKR^{Bio} prozedurales fachdidaktisches Wissen valide messen kann. Der validierte SKR^{Bio} wird in Studie 3 in Kombination mit einem Fragebogen zur Evolution eingesetzt, um erste Erkenntnisse über die Zusammenhänge zwischen deklarativem und prozeduralem Professionswissen bei Lehramtsstudierenden zu generieren. Hierbei zeigten sich Zusammenhänge zwischen beiden Wissenstypen innerhalb des Professionswissens.

Die vorliegende Arbeit stellt eine neue simulierte Klassenraumumgebung vor, welche die Erfassung prozeduraler Facetten des Professionswissens ermöglicht. Mithilfe des SKR^{Bio} konnten bereits theoretisch und empirisch beschriebene Schwierigkeiten im Umgang mit spezifischen Fehlvorstellungskategorien repliziert werden. Außerdem lieferte der SKR^{Bio} erste Hinweise auf den Zusammenhang zwischen deklarativem und prozeduralem biologiebezogenen Professionswissen, wodurch kognitionspsychologische Annahmen zur Interaktion beider Wissenstypen bestätigt wurden.

SUMMARY

Professional knowledge is considered the core of professional competence and is the central prerequisite for a quality teaching as well as the associated learning success of the students. In general, professional knowledge can be differentiated into knowledge areas – content, pedagogical content and pedagogical knowledge – and into knowledge types – declarative and procedural knowledge. Previous research has conducted numerous studies to systematically investigate teachers' professional knowledge. There has been a lack of instruments that can capture not only declarative but also procedural professional knowledge. One solution is to use simulated classroom environments in which prospective teachers can apply their declarative knowledge they have gained in academic discourse. In the biology classroom, procedural knowledge becomes particularly relevant when assessing shown student achievement. Especially in the context of biological evolution, which is the integrative framework of biology education, misconceptions are often expressed to explain evolutionary processes.

This dissertation includes three empirical studies. All studies investigate facets of procedural professional knowledge with a focus on biology pedagogical content knowledge. Study 1 uses a digital instrument – the Student Inventory – to investigate the extent to which virtual student exams can be assessed on the basis of pre-service teachers' professional knowledge. Here, deficits in procedural pedagogical content knowledge as well as pedagogical knowledge were revealed. In study 2, a simulated classroom environment – the Simulated Classroom Biology (SCR^{Bio}) – that was further developed in the context of this dissertation was piloted and validated. The results indicated that the SCR^{Bio} can validly measure procedural pedagogical content knowledge. The validated SCR^{Bio} is used in Study 3 in combination with a questionnaire on evolution to generate initial findings on the interrelationships between declarative and procedural professional knowledge of pre-service teachers. Here, correlations between both types of knowledge within professional knowledge were found.

This dissertation introduces a new simulated classroom environment that allows the acquisition of procedural facets of professional knowledge. Using the SCR^{Bio}, previously theoretically and empirically described difficulties in dealing with specific misconception categories could be replicated. In addition, the SCR^{Bio} provided initial evidence of the interrelationship between declarative and procedural biology-related professional knowledge, confirming cognitive psychological assumptions about the interaction of the two types of knowledge.

INHALTSVERZEICHNIS

ZUSAMMENFASSUNG	I
SUMMARY	III
INHALTSVERZEICHNIS	V
LISTE DER ABBILDUNGEN	IX
LISTE DER TABELLEN	X
1 EINLEITUNG	1
2 THEORETISCHER HINTERGRUND UND AKTUELLER STAND DER FORSCHUNG	3
2.1 Professionswissen als Kern professioneller Kompetenz	3
2.1.1 Struktur und Konzeptualisierung von Professionswissen in Wissensbereiche	4
2.1.2 Struktur und Konzeptualisierung von Professionswissen in Wissenstypen .	12
2.2 Die biologische Evolution und Evolutionstheorie	18
2.2.1 Lehren und Lernen der Evolution	18
2.2.2 Fehlvorstellungen zur natürlichen Selektion	20
2.3 Simulierte Klassenraumumgebungen	22
2.3.1 Der Simulierte Klassenraum	24
3 ZIELSETZUNG DER ARBEIT UND ZUSAMMENFASSUNG DER DURCHGEFÜHRTEN STUDIEN	25
3.1 Studie 1 (Kapitel 4): Measuring Biology Trainee Teachers' Professional Knowledge about Evolution – Introducing the Student Inventory	26
3.2 Studie 2 (Kapitel 5): Simulated Classroom Biology (SCR^{Bio}) - a Simulated Classroom Environment for Assessing the Action-Oriented Professional Knowledge of Pre-Service Teachers about Evolution.....	26
3.3 Studie 3 (Kapitel 6): Der Simulierte Klassenraum Biologie: Erfassung deklarativen und prozeduralen Wissens bei Lehramtsstudierenden der Biologie	27

4	STUDIE 1: MEASURING BIOLOGY TRAINEE TEACHERS' PROFESSIONAL KNOWLEDGE ABOUT EVOLUTION – INTRODUCING THE STUDENT INVENTORY	31
4.1	Introduction	33
4.2	Theoretical Background	34
4.2.1	Teaching and Learning Evolution	34
4.2.2	Biology Teachers' Professional Knowledge	38
4.3	Research Questions	40
4.4	Method	42
4.4.1	Sample	42
4.4.2	The Student Inventory (SI)	43
4.4.3	Assessment of Declarative and Procedural Knowledge with the SI	45
4.4.4	Procedure	46
4.4.5	Analyses.....	47
4.5	Results	47
4.6	Discussion.....	49
4.6.1	Limitations.....	51
4.6.2	Implications for Further Research and Practical Implications	52
4.7	References	53
4.8	Supplemental Material	64
4.8.1	Overview About the Student Written Answers	64
4.8.2	Examples of Items in the Questionnaire on Evolution	67
4.8.3	Screenshots of Virtual Student Exams in the SI.....	68
5	STUDIE 2: THE SIMULATED CLASSROOM BIOLOGY (SCR^{BI0}) – A SIMULATED CLASSROOM ENVIRONMENT FOR CAPTURING THE ACTION-ORIENTED PROFESSIONAL KNOWLEDGE OF PRE-SERVICE TEACHERS ABOUT EVOLUTION	73
5.1	Introduction	74
5.2	Theoretical Background	75
5.2.1	Teachers' Professional Knowledge	75
5.2.2	Simulated Classroom Environments	77

5.2.3 Teaching and Learning Evolution	78
5.3 Research Questions	79
5.4 Methods.....	80
5.4.1 Sample	80
5.4.2 Development of the Simulated Classroom Biology (SCR ^{Bio})	81
5.4.3 Procedure	82
5.4.4 Validity	85
5.5 Results	86
5.6 Discussion.....	89
5.6.1 Limitations	90
5.6.2 Implications for Further Research	92
5.7 References	92
6 STUDIE 3: DER SIMULIERTE KLASSENRAUM BIOLOGIE – ERFASSUNG DEKLARATIVEN UND PROZEDURALEN WISSENS BEI LEHRAMTSSTUDIERENDEN DER BIOLOGIE?	103
6.1 Einleitung.....	106
6.2 Theoretischer Hintergrund	108
6.2.1 Professionswissen von Biologielehrkräften.....	108
6.2.2 Lehren und Lernen der biologischen Evolution	110
6.2.3 Fehlvorstellungen zum Prozess der natürlichen Selektion	111
6.3 Fragestellung und Hypothesen	112
6.4 Methodisches Vorgehen.....	113
6.4.1 Stichprobe und Ablauf.....	113
6.4.3 Testinstrumente	113
6.4.4 Operationalisierung von deklarativem und prozeduralem Wissen.....	116
6.4.5 Ablauf	116
6.4.6 Datenauswertung	117
6.5 Ergebnisse	117
6.5.1 Deklaratives Wissen	117
6.5.2 Prozedurales Wissen.....	118
6.5.3 Zusammenhang deklarativen und prozeduralen Wissens.....	118

6.6 Diskussion	120
6.6.1 Gezeigtes prozedurales Wissen im SKR ^{Bio}	121
6.6.2 Zusammenhang zwischen deklarativem und prozeduralem Wissen	122
6.7 Fazit und Ausblick	124
6.8 Literatur	126
7 ZUSAMMENFASSUNG DER ERGEBNISSE DER DREI STUDIEN	139
7.1 Studie 1: Measuring Biology Trainee Teachers' Professional Knowledge about Evolution – Introducing the Student Inventory	139
7.2 Studie 2: Simulated Classroom Biology (SCR ^{Bio}) – a Simulated Classroom Environment for Assessing the Action-Oriented Professional Knowledge of Pre-Service Teachers About Evolution	140
7.3 Studie 3: Der Simulierte Klassenraum – Erfassung deklarativen und prozeduralen Wissens bei Lehramtsstudierenden der Biologie.....	141
8 DISKUSSION UND AUSBLICK	143
8.1 Übergreifende Diskussion.....	143
8.1.1 Entwicklung des SKR ^{Bio}	143
8.1.2 Operationalisierung und Differenzierung deklarativen und prozeduralen Wissens.....	145
8.1.3 Erfassung des prozeduralen FDW mit dem SKR ^{Bio}	147
8.2 Grenzen der Studien	148
8.3 Implikationen für zukünftige Forschung und universitäre Praxis	150
8.4 Ausblick.....	151
LITERATURVERZEICHNIS.....	153
ERKLÄRUNG.....	173
DANKSAGUNG	175

LISTE DER ABBILDUNGEN

Abbildung 2.1 COACTIV-Modell zur professionellen Handlungskompetenz und zum Professionswissens (verändert nach Baumert und Kunter, 2006).	6
Abbildung 2.2 Magnusson-Modell zum fachdidaktischen Wissen (verändert nach Magnusson et al., 1999).	9
Abbildung 2.3 Refined Consensus Modell of PCK (verändert nach Carlson und Daehler, 2019).	17
Abbildung 4.1 An overview of the investigated knowledge types, knowledge domains, and knowledge facets, as well as the operationalization within the SI and related research questions.	42
Abbildung 4.2 Main effect of the previous performance in the multiple-choice test.	48
Abbildung 4.3 Screenshot 1 of the SI.....	69
Abbildung 4.4 Screenshot 2 of the SI.....	70
Abbildung 4.5 Screenshot 3 of the SI.....	71
Abbildung 4.6 Screenshot 4 of the SI.....	72
Abbildung 5.1 Screenshot of the teaching sequence in the SCR ^{Bio}	84
Abbildung 5.2 Control group: Relative frequency of diagnosis of students' answers with regard to expressed misconceptions categories or scientific reasoning.	87
Abbildung 5.3 Experimental group: Relative frequency of diagnosis of students' answers with regard to expressed misconceptions categories or scientific reasoning.....	88
115	
Abbildung 6.1 Darstellung des SKR ^{Bio} während der Fragensauswahl durch die Lehramtsstudierenden.	115

LISTE DER TABELLEN

Tabelle 3.1 Übersicht über die durchgeführten Studien.....	28
Tabelle 3.1 Fortsetzung	29
Tabelle 3.1 Fortsetzung	30
Tabelle 4.1 Scheme of the variation of qualities of students' written answers and of the students' performance in the multiple-choice test in the SI.....	45
Tabelle 5.1 Scheme for the development of evolutionary biological questions and possible student answers in the SCR ^{Bio}	81
Tabelle 5.2 Example question and possible student answers about the red bentgrass (botanical organism, gain of a trait, physiological trait)	82
Tabelle 5.3 Results of the ranking, level, and differentiation component subdivided by control and experimental group.	89
Tabelle 6.1 Produkt-Moment-Korrelationen zwischen dem gezeigten deklarativen Wissen im Fragebogen und dem gezeigten prozeduralen Wissen im SKR ^{Bio} der Lehramtstudierenden.....	120

1 EINLEITUNG

Die universitäre Ausbildung zukünftiger Lehrerinnen und Lehrer wird als maßgeblicher Faktor der Qualitätssicherung des naturwissenschaftlichen Unterrichts angesehen (Europäische Kommission, 2015). Hierbei legt die Entwicklung eines fundierten Professionswissens die Grundlage für angehende Lehrkräfte, den Herausforderungen des realen Unterrichtens gewachsen zu sein (z. B. Baumert und Kunter, 2006; Kirschner, 2013; Sadler et al., 2013). Im Zuge von Forschungsarbeiten zur Beschreibung des Professionswissens kristallisierten sich drei zentrale Wissensbereiche – das fachliche, fachdidaktische und pädagogische Wissen – heraus (z. B. Berry et al., 2015; Cochran et al., 1993; Gess-Newsome und Lederman, 1999; Grossman, 1990; Loughran et al., 2006; Shulman, 1986). Zur Konzeptualisierung des Professionswissens hat sich in der deutschsprachigen Wissenschaftslandschaft das aus der Bildungswissenschaft stammende *COACTIV*-Modell durchgesetzt, welches das fachliche, fachdidaktische und pädagogische Wissen weiter in spezifische Wissensfacetten differenziert (Baumert und Kunter, 2006). In Anlehnung an psychologische Ansätze hat sich in der Forschung zum Professionswissen eine dichotome Einteilung in deklaratives („Wissen, dass“) und prozedurales Wissen („Wissen, wie“) etabliert (z. B. Baumert und Kunter, 2013; Blömeke et al., 2010; König et al., 2014). Hierbei markiert das deklarative Wissen der zukünftigen Lehrerinnen und Lehrer den Startpunkt für die Entwicklung prozeduralen Wissens (z. B. Anderson, 1983; Shavelson et al., 2005). Innerhalb der Lehrerprofessionalisierungsforschung wurden zahlreiche Large-Scale-Studien durchgeführt, welche vor allem die Messung des deklarativen Wissens via Fragebögen fokussierten (z. B. Baumert und Kunter, 2013; Blömeke et al., 2008; Kleickmann et al., 2014; Tepner et al., 2012), wohingegen die Erfassung des prozeduralen Wissens weitestgehend vernachlässigt wurde.

Die vorliegende Dissertation beschäftigt sich mit der Untersuchung des Zusammenhangs von deklarativen und prozeduralen Facetten des Professionswissens angehender Biologielehrkräfte zur Evolution. Die Evolutionstheorie gilt als der integrative Rahmen des Biologieunterrichts und ist die Grundlage aller biologischen Themen und eines konzeptuellen Verständnisses der Biologie (z. B. Basel et al., 2014; Bishop und Anderson, 1990; Furtak, 2012; Zabel und Gropengießer, 2011). In Studien zum Lehren und Lernen der Evolution werden zahlreiche Probleme beschrieben, welche vor allem auf einem fehlerhaften Verständnis des zentralen Prozesses der Evolution – der natürlichen Selektion – beruhen (z. B. BouJaoude et al., 2011; Gregory, 2009; Großschedl et al., 2014; Lucero et al., 2019; Nehm und Schonfeld, 2007). Die angehenden Lehrerinnen und Lehrer sind der entscheidende Faktor, fehlerhafte Denkweisen der zukünftigen Schülerinnen und Schüler frühzeitig zu erkennen und unterstützende Maßnahmen einzuleiten, damit konzeptuelles

EINLEITUNG

evolutionsbiologisches Wissen erworben werden kann (Gregory, 2009). Hierfür wird ein elaboriertes biologiebezogenes Professionswissen notwendig, welches im akademischen Kontext erworben wird und sich positiv auf die Unterrichtsqualität sowie auf das Lernen der Schülerinnen und Schüler auswirkt (z. B. Baumert et al., 2010; Hill et al., 2005; Park et al., 2011; Roth et al., 2011). Um die Facetten des biologischen Professionswissens angehender Lehrkräfte in einer handlungsnahen Anwendungssituation beobachten und messen zu können, bieten sich simulierte Klassenraumumgebungen an. Innerhalb dieser Simulationen können die angehenden Lehrerinnen und Lehrer ihr professionelles Wissen interaktiv einbringen, ohne dabei reale Konsequenzen ihres Handelns zu erfahren (z. B. van Berkum und de Jong, 1991; Hixon und So, 2009; Plass et al., 2009). Gleichzeitig ermöglicht die experimentelle Kontrolle der Variablen innerhalb einer simulierten Klassenraumumgebung, differenzierte Einblicke über die Beschaffenheit und Struktur des prozeduralen professionellen Wissens zu generieren (Badiee und Kaufmann, 2015).

Im einleitenden Kapitel dieser Arbeit wird der theoretische Hintergrund und aktuelle Stand der Forschung erläutert. Anfänglich wird die Struktur und Konzeptualisierung des Professionswissens beschrieben (Kapitel 2.1), wobei explizit auf die Unterscheidung in Wissensbereiche (Kapitel 2.1.1) und Wissenstypen (Kapitel 2.1.2) eingegangen wird. Im Anschluss folgt ein Kapitel zur biologischen Evolution und Evolutionstheorie (Kapitel 2.2), in welchem vorerst die Herausforderungen des Lehrens und Lernens der Evolution (Kapitel 2.1.1) thematisiert wird, um aufbauend darauf spezifische Fehlvorstellungen zum Prozess der natürlichen Selektion (Kapitel 2.1.2) zu beschreiben. Im letzten theoretischen Kapitel werden simulierte Klassenraumumgebungen eingeführt (Kapitel 2.3) und darauf folgend der Simulierte Klassenraum vorgestellt, welcher im Rahmen dieser Dissertation für die Beantwortung biologiedidaktischer Fragestellungen weiterentwickelt wurde (Kapitel 2.3.1). Anschließend werden die Ziele und dazugehörigen Fragestellungen dieser Dissertation erarbeitet sowie ein kurzer Überblick über die durchgeführten Studien gegeben (Kapitel 3). Nach der Vorstellung der drei im Rahmen der Doktorarbeit durchgeführten empirischen Studien (Kapitel 4 bis 6) werden deren Ergebnisse zusammengefasst (Kapitel 7) und studienübergreifend sowie in Hinblick auf die Implikationen für zukünftige Forschung und universitäre Praxis diskutiert (Kapitel 8).

2 THEORETISCHER HINTERGRUND UND AKTUELLER STAND DER FORSCHUNG

2.1 Professionswissen als Kern professioneller Kompetenz

Die professionelle Kompetenz von Lehrerinnen und Lehrern gilt als wesentliche Voraussetzung für einen erfolgreichen Unterricht und den daraus resultierenden Schülerleistungen (z. B. Baumert und Kunter, 2006; Besser und Krauss, 2009; Hattie, 2009). Hinsichtlich der Kompetenzmodellierung liegen mehrere, zum Teil empirische Forschungsarbeiten vor, welche unterschiedliche Aspekte in die Konzeptualisierung der professionellen Kompetenz einbeziehen (z. B. Baumert und Kunter, 2006; Klieme und Leutner, 2006; Weinert, 2001a). Im Gegensatz zur Intelligenz, welche kaum durch äußere Faktoren beeinflusst werden kann, sind Kompetenzen erlernbar und können während der akademischen Laufbahn erworben werden (Weinert, 2001b). Der Kompetenzbegriff nach Weinert (2001a,b) beinhaltet neben kognitiven Voraussetzungen auch motivationale, volitionale sowie soziale Bereitschaften und Fähigkeiten, die eine Lehrperson in die Lage versetzen, in spezifischen unterrichtlichen Anforderungskontexten adäquat zu reagieren. Hierbei wird vor allem die Handlungskompetenz betont, welche durch das dynamische Zusammenwirken der zuvor genannten Aspekte definiert wird (Weinert, 2001a). Übertragen auf die professionelle Kompetenz betrifft die Handlungskompetenz den beruflichen Alltag von Lehrerinnen und Lehrern, beispielsweise bei der Stundenplanung oder dem Unterrichten. Im deutschsprachigen Raum prägte vor allem die *COACTIV*-Studie die Konzeptualisierung zur professionellen Kompetenz von Lehrkräften, in welcher kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz von Schülerinnen und Schülern untersucht wurde (Kunter et al., 2011). Basierend auf den Überlegungen von Shulman (1986, 1987) resultierte ein häufig zitiertes Modell der professionellen Kompetenz von Lehrkräften, welches zwischen Professionswissen, Überzeugungen / Werthaltungen, motivationalen Orientierungen und selbstregulativen Fähigkeiten differenziert (Baumert und Kunter, 2006). Hierbei handelt es sich nicht um ein hierarchisches Modell, da alle Kompetenzaspekte miteinander interagieren (Baumert und Kunter, 2013). Innerhalb der Überzeugungen geht es um individuelle Vorstellungen sowie Annahmen von Lehrerinnen und Lehrern mit einer wertenden Komponente, die sich beispielsweise auf den Lehr-Lern-Kontext (z. B. spezifische Unterrichtsansätze und -methoden), auf das Bildungssystem (z. B. Bildungsreformen) oder auf die Gesellschaft (z. B. kulturelle Normen und Werte) beziehen (Kunter und Pohlmann, 2009; Woolfolk et al., 2006). Motivationale Orientierungen basieren auf der erwarteten Selbstwirksamkeit, den intrinsischen Orientierungen sowie den

Zielvorstellungen der Lehrkraft (Richardson et al., 2014). Die Selbstregulation bezieht sich vor allem auf die empfundene Belastbarkeit von Lehrkräften, welche je nach beruflichen Umweltfaktoren (z. B. Arbeitsbelastung, Klassengröße) positive oder negative Auswirkungen auf das unterrichtliche Handeln haben können (Kunter und Pohlmann, 2009). Als Kern professioneller Kompetenz wird das Professionswissen beschrieben, welches als berufsspezifisches Wissen innerhalb der universitären Ausbildung vermittelt und in der Praxis vertieft werden kann (z. B. Großschedl, 2015a; Kleickmann et al., 2014; Mahler et al., 2017). Dem Professionswissen kommt eine besondere Bedeutung zu, da es als wesentliche Voraussetzung für guten Unterricht gilt und somit zu besseren Schülerleistungen führt (z. B. Baumert und Kunter, 2006; Borowski et al., 2010; Neubrand, 2006; Sadler et al., 2013).

In den folgenden Kapiteln werden die Struktur sowie Konzeptualisierungen von Professionswissen vorgestellt. Es werden zunächst unterschiedliche Modelle des Professionswissens beschrieben, um darauf aufbauend die verschiedenen Wissensbereiche zu erörtern. Hierbei liegt ein starker Fokus auf der Entwicklung, der Struktur und den spezifischen Facetten des fachdidaktischen Wissens (Kapitel 2.1.1). Anschließend wird beschrieben, inwiefern sich das Professionswissen in Anlehnung an kognitionspsychologische Ansätze weiter in verschiedene Wissenstypen differenzieren lässt und welche Konzeptualisierungen dieser Wissenstypen in naturwissenschafts- und mathematikdidaktischen Studien Anwendung gefunden haben (Kapitel 2.1.2).

2.1.1 Struktur und Konzeptualisierung von Professionswissen in Wissensbereiche

Im Unterricht nimmt die Lehrkraft eine wesentliche Rolle ein, denn sie ist der entscheidende Faktor für ein motivierendes Lernumfeld sowie für die Leistungsentwicklung der Schülerinnen und Schüler (z. B. Hill et al., 2005; Baumert und Kunter, 2011b; Sadler et al., 2013). Unstrittig ist, dass die Kompetenzen von Lehrkräften, darunter auch das Professionswissen, eine zentrale Voraussetzung für erfolgreichen Unterricht sind (Hattie, 2005; 2009). Das Professionswissen bezieht sich auf das Wissen, welches eine Lehrkraft benötigt, um den Anforderungen des Berufs gerecht zu werden (Baumert et al., 2010; Kirschner, 2013). Hierbei ist Professionswissen domänenspezifisch, gut organisiert und hierarchisch strukturiert, um in unterschiedlichsten Handlungskontexten angewendet zu werden (Baumert und Kunter, 2013).

Eine erste Auflistung von Aspekten, welche das Professionswissen von Lehrkräften umfassen sollte, kommt von Shulman (1986). Er differenziert anfänglich in fachinhaltsbezogenes Wissen, fachdidaktisches Wissen, pädagogisches Wissen und in Wissen über das Curriculum. Besonders relevant sind hierbei die ersten beiden Wissensbereiche, da diese in der Bildungs- sowie fachdidaktischen Forschung lange vernachlässigt wurden. Das fachinhaltbezogene Wissen adressiert nach Shulman (1986)

die Kenntnis von fachlichen Fakten und deren Vernetzung zu Konzepten sowie das Verständnis darüber, wie dieses Wissen strukturiert ist. Weiterhin misst Shulman (1986) dem Begründungswissen innerhalb des Fachwissens eine große Bedeutung zu, denn dieses wird besonders dann wichtig, wenn es um die Vermittlung von Werten und Anwendungsbereichen des Wissens geht. Das fachdidaktische Wissen beschreibt Shulman (1986) als das Wissen, welches eine Lehrkraft benötigt, um unterrichten zu können. Es umfasst dabei das Wissen über typische Lernschwierigkeiten, zum Beispiel nicht-wissenschaftliche Denkweisen von Schülerinnen und Schülern, sowie Instruktionsstrategien, um diesen Lernschwierigkeiten adäquat zu begegnen. Weiterhin ist das Wissen zu unterschiedlichen Repräsentationsformen ein wichtiger Bestandteil des fachdidaktischen Wissens. Fachliches und fachdidaktisches Wissen sind nach Shulman (1986) eng miteinander verbunden. Später ergänzte Shulman (1987) zu der bereits eingeführten Differenzierung des Professionswissens das Wissen über Lernstrategien, das Wissen über Bildungsinhalte und das Wissen über die philosophischen und historischen Ziele von Bildung. In der bisherigen Forschung wurde sich vor allem auf die im Eingang beschriebenen drei Wissensbereiche – fachliches Wissen (FW), fachdidaktisches Wissen (FDW) und pädagogisches Wissen (PW) – konzentriert (z. B. Baumert et al., 2010; Bromme, 1994; Förtsch et al., 2018; Grossman, 1990; Großschedl et al., 2015a; Jüttner et al., 2013; Magnusson et al., 1999; Mahler et al., 2017; Sadler et al., 2013).

In der Professionsforschung von Lehrkräften haben sich zahlreiche Projekte mit der Modellierung von Professionswissen in den naturwissenschaftlichen Fächern beschäftigt (z. B. Borowski et al., 2010, 2011; Kröger et al., 2013; Lange et al., 2012; Olszewski, 2010a; Riese und Reinhold, 2010). Ein in der Forschung vielfach betrachtetes und akzeptiertes Modell zur Konzeptualisierung von Professionswissen liefern Baumert und Kunter (2006; s. Abbildung 2.1). Dieses Modell beinhaltet kognitive sowie affektiv-motivationale Aspekte und diente als Basis des Projekts *COACTIV* (Kunter et al., 2011). Innerhalb dieses Modells wird das Professionswissen, neben Organisations- und Beratungswissen, in FW, FDW sowie PW differenziert und war Ausgangspunkt vieler weiterer Studien in der deutschsprachigen Lehrerprofessionsforschung zu naturwissenschaftlichen Fächern (z. B. Förtsch et al., 2018; Großschedl et al., 2015a; Jüttner et al., 2013; Riese und Reinhold, 2012). Trotzdem gibt es in den Bildungswissenschaften sowie in der naturwissenschaftlichen Fachdidaktik unterschiedliche Ansichten darüber, inwiefern diese Wissensbereiche angemessen beschrieben und in Wissensfacetten strukturiert werden können.

Angelehnt an das *COACTIV*-Modell und die zuvor beschriebenen Wissensbereichen innerhalb des Professionswissens konzentriere ich mich in meiner

THEORETISCHER HINTERGRUND

Dissertation auf die Beschreibung und Untersuchung von Facetten des fachlichen, fachdidaktischen und pädagogischen Wissens.

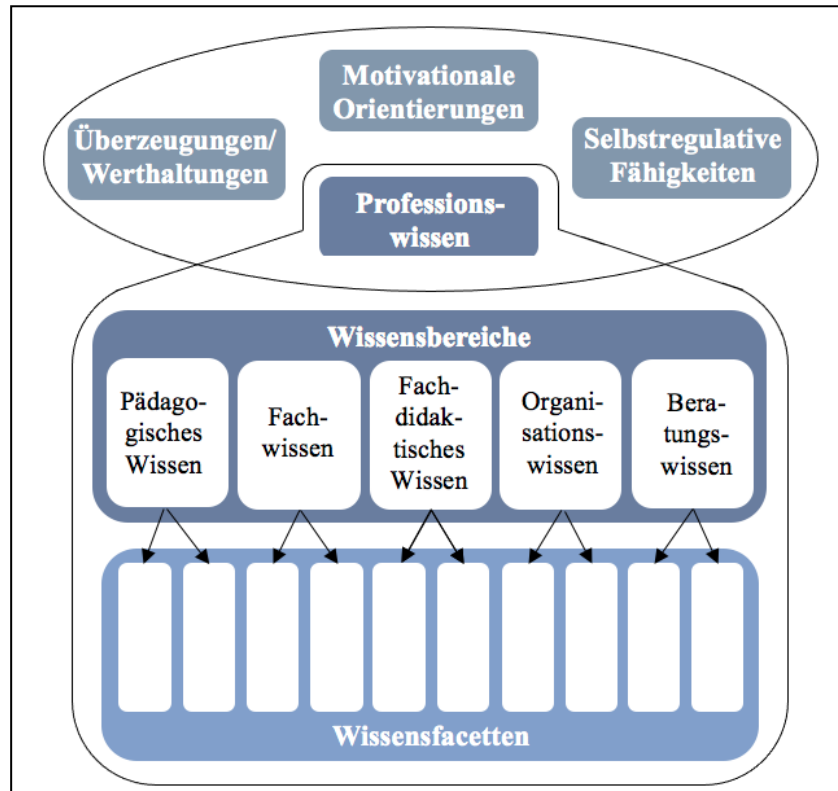


Abbildung 2.1 COACTIV-Modell zur professionellen Handlungskompetenz und zum Professionswissen (verändert nach Baumert und Kunter, 2006).

FW. Das FW gilt als Grundvoraussetzung für einen erfolgreichen und effektiven Fachunterricht (z. B. Ball et al., 2005; Shulman, 1986, 1987; Jüttner und Neuhaus, 2013). Definiert wird FW als das vertiefte Schulwissen sowie als das im akademischen Diskurs erlangte FW (z. B. Baumert und Kunter, 2006; Riese und Reinhold, 2009). Im Allgemeinen bezieht sich FW auf das Wissen über Fakten und Begriffe sowie das konzeptuelle Verständnis der zu unterrichtenden Inhalte (Shulman, 1986). Mit konzeptuellem Verständnis ist die Identifikation und Organisation zentraler Konzepte und Prinzipien innerhalb eines Inhaltsbereichs gemeint (Shulman, 1986). Nach Bromme (1992) wird FW durch die Philosophie des Fachinhalts ergänzt, um den Nutzen sowie das Verhältnis dieses Fachinhalts zu anderen Wissen- und Lebensbereichen darzustellen. Innerhalb des FW eines Faches wird eine Domänenspezifität angenommen, wonach ein elaboriertes Wissen in einem Themengebiet nicht gleichzeitig ein hohes Wissen in einem anderen Themengebiet voraussetzt (Hill et al., 2004). Zur Binnendifferenzierung des FW haben Ball et al. (2005) unterschiedliche Niveaustufen vorgeschlagen, das allgemeine fachliche Wissen, welches von gut ausgebildeten Erwachsenen erwartet wird und das Wissen zur expliziten Ausübung des Lehrberufs. In darauf folgenden Studien werden weitere Niveaustufen modelliert, um

das FW nach dem Grad der Ausprägung zu kategorisieren (z. B. Baumert et al., 2010; Blömeke et al., 2008, 2009; Riese und Reinhold, 2009, 2010; Tatto et al., 2008). Im Kontext unterschiedlicher Ansätze zur Konzeptualisierung von FW herrscht Einigkeit darüber, dass zur angemessenen Vorbereitung und erfolgreichen Durchführung von Unterricht mehr als nur Schulwissen, d.h. ein tieferes Verständnis des Fachinhalts, erforderlich ist (z. B. Blömeke et al., 2009; Krauss et al., 2008).

Studien zur Messung biologiebezogenen Professionswissens benannten unterschiedliche Themenkomplexe, auf deren Basis das FW angehender Biologielehrkräfte erfasst werden sollte. So differenziert Jüttner und Neuhaus (2013) das FW in Wissen zur Zytologie und Neurobiologie sowie zu Wirbeltieren und Pflanzen, wohingegen Großschedl et al. (2015a) die Studieninhalte Ökologie, Physiologie, Morphologie, Genetik und Molekularbiologie sowie Evolution auswählen. Nach Shulmans (1986) Definition des konzeptuellen Wissens könnte beispielsweise das FW zur Evolution weiter in die Evolutionsmechanismen (z. B. Selektion, Artbildung, Koevolution), die Stammesgeschichte und ihre Rekonstruktion (z. B. Stammbäume, Homologie/Analogie, Fossilien, rudimentäre Organe) sowie in divergente Evolutionstheorien (z. B. Curvier, Lamarck, Wallace, Darwin) differenziert werden (Großschedl et al., 2015b). Dieses tiefere Verständnis innerhalb eines Themenbereichs ist notwendig, da die Unterstützung der Schülerinnen und Schüler beim Lernen fachlichen Wissens mehr benötigt, als die bloße Vermittlung von isolierten Fakten und Begriffen (Shulman, 1986).

PW. Das pädagogische bzw. pädagogisch-psychologische Wissen von Lehrerinnen und Lehrern bezieht sich nach Shulman (1986) auf den konkreten Unterrichtskontext im Klassenverband und umfasst die Facetten der Klassenführung und der Klassenorganisation. Ergänzt und spezifiziert werden die von Shulman (1986) eingeführten Facetten durch Reynolds (1992), welcher das Wissen über Strategien und Mittel zur Herstellung lernförderlicher Rahmenbedingungen, das Wissen zur Bedeutung echter Lernzeit und das Wissen über Unterrichtsregeln hinzufügt. Renkl (2008) formuliert im Sinne der effektiven Klassenführung und angelehnt an Reynolds (1992) Ideen sechs weitere Prinzipien, die sich beispielsweise auf die Verhinderung von ungenutzter Unterrichtszeit, Etablierung eines effizienten Regelsystems, Vermeidung von Unterrichtsstörungen oder Integration eines angemessenen Anforderungsniveaus beziehen. Im *COACTIV*-Projekt (Baumert und Kunter, 2006) wird eine allgemeine Konzeptualisierung von PW präsentiert, welche zwischen konzeptuellem bildungswissenschaftlichen Grundlagenwissen, allgemeindidaktischem Konzeptions- und Planungswissen, Wissen über Unterrichtsführung und Orchestrierung von Lerngelegenheiten und Wissen über fächerübergreifende Prinzipien des Diagnostizierens, Prüfens und Bewertens unterscheiden. In der Nachfolgestudie *COACTIV-R* (Voss und Kunter, 2011) wird das PW weiter spezifiziert und die Facetten Klassenführung, Unterrichtsmethoden,

THEORETISCHER HINTERGRUND

Leistungsbeurteilung sowie individuelle Lernprozesse eingeführt. Unter Klassenführung werden, angelehnt an Helmke (2009), der Umgang mit Disziplinproblemen sowie die Regeltransparenz und das Vorbeugen von Störungen verstanden. Um Unterricht erfolgreich gestalten zu können, ist das Wissen über die Vielfalt von Unterrichtsmethoden hilfreich (Helmke, 2009). Dabei steigt die Qualität von Unterricht nicht zwangsläufig durch die Verwendung zahlreicher Unterrichtsmethoden, sondern vielmehr durch die optimale Passung von Unterrichtsmethode und dem jeweiligen Unterrichtsziel. Das Wissen zur Unterstützung individueller Lernprozesse umfasst vor allem effektive Strategien, um ein erfolgreiches Lernen der Schülerinnen und Schüler zu gewährleisten. Die Leistungsbeurteilung fokussiert die formativen (innerhalb einer Unterrichtsstunde) und summativen Rückmeldungen (am Ende einer Unterrichtsstunde). Hierbei geht es um fächerübergreifende Aspekte der Leistungsrückmeldung, wie zum Beispiel zum Zweck der Optimierung der folgenden Unterrichtssequenz oder im Rahmen der Notenvergabe. Deutlich abgegrenzt von der Leistungsbeurteilung im pädagogischen Sinne werden die Beurteilungen im fachlichen (z. B. wissenschaftliche Korrektheit) und fachdidaktischen Kontext (z. B. zugrundeliegender Schülerfehlvorstellungen; Tepner et al., 2012).

FDW. Mitte der 80er Jahre definierte Shulman (1986) erstmals den Begriff des FDW, welcher als lehrerspezifischer Wissensbereich ein Amalgam aus FW und PW darstellt. Somit spannt das FDW eine weitere Wissensdomäne innerhalb des lehrerbezogenen Professionswissens auf, welche die Lehrerin oder den Lehrer befähigen, Unterrichtsinhalte schülerzentriert zu vermitteln, und auf eventuelle Probleme angemessen reagieren zu können (Shulman, 1986). Die Einführung des FDW durch Shulman während seines Vortrags auf der Tagung der *American Educational Research Association* im Jahr 1985 fiel in eine Zeit, in der eine große Unzufriedenheit mit dem Bildungssystem sowie der Lehrerausbildung in den USA herrschte (Carlsen, 1999). Gleichzeitig markierte die vorgestellte Konzeptualisierung des FDW den Startpunkt für eine Extensivierung der Forschung über das professionelle Wissen von Lehrkräften (Gess-Newsome, 1999).

In Shulmans (1986) initialer Beschreibung des FDW werden zwei Facetten definiert: das Wissen, welches Lehrkräften ermöglicht, den Lehrstoff verständlich für die Schülerinnen und Schüler darzustellen sowie das Wissen über typische Lernschwierigkeiten und Vorstellungen bzw. Denkweisen von Schülerinnen und Schülern zu einem spezifischen Unterrichtsthema. Demnach ist das FDW eng mit dem FW verbunden (Shulman, 1986). Nach den ersten Arbeiten Shulmans (1986, 1987) haben sich zahlreiche Wissenschaftlerinnen und Wissenschaftler der Erforschung des FDW angeschlossen. In der Konzeptualisierung nach Grossman (1990) umfasst das FDW neben dem fachbezogenen Wissen über Schülerkognition und Instruktionsstrategien noch das Wissen über Lehrpläne. Später ergänzten Hashweh (2005) und Magnusson et al. (1999) das Wissen über Beurteilungen (engl. *Assessment* genannt). Es wird deutlich, dass in der

Wissenschaft unterschiedliche Vorstellungen darüber vorliegen, welche Facetten dem FDW zugeordnet werden können. Einen Überblick über die Vielzahl der Konzeptualisierungen bietet der Literaturüberblick von Park und Oliver (2008), welche ihre Analysen auf Basis von zwölf Veröffentlichungen zum FDW durchführten. Dabei wird deutlich, dass vor allem die Facetten: Wissen über Instruktionsstrategien und das Wissen über Schülervorstellungen als Kern des FDW angesehen werden (z. B. Grossman, 1990; Hashweh, 2005; Loughran et al., 2006; Magnusson et al., 1999; Shulman, 1987; Tamir, 1988). Im Folgenden werden die Kernfacetten vom FDW am häufig zitierten Modell (s. Abbildung 2.2) nach Magnusson et al. (1999) beschrieben und exemplarisch auf das Thema Evolution im Biologieunterricht bezogen.

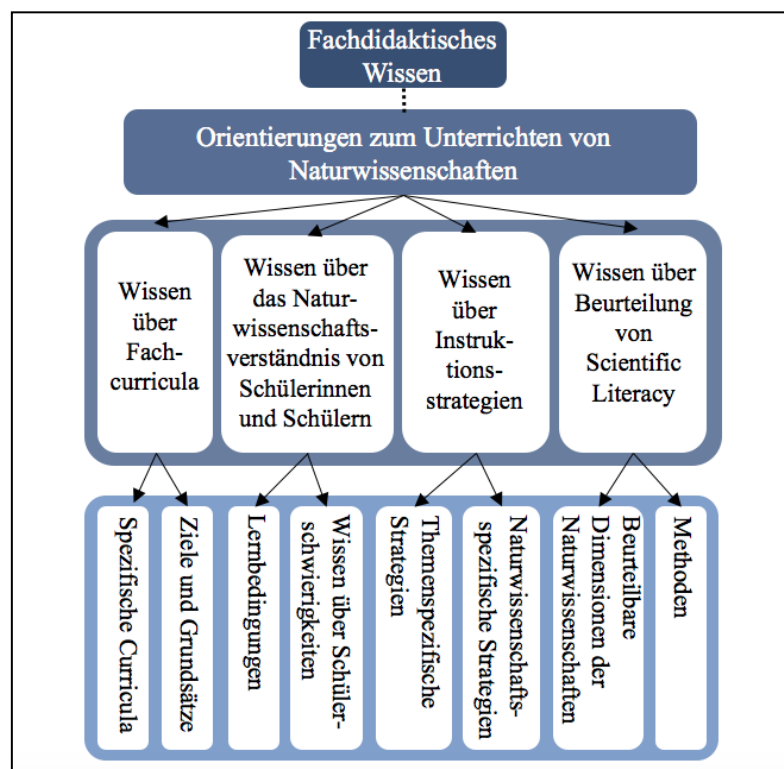


Abbildung 2.2 Magnusson-Modell zum fachdidaktischen Wissen (verändert nach Magnusson et al., 1999).

Das Wissen über Instruktionsstrategien umfasst neben dem Wissen über fachspezifische auch themenspezifische Instruktionsstrategien (Magnusson et al., 1999). Hierbei sind nach Magnusson et al. (1999) fachspezifische Instruktionsstrategien als naturwissenschaftsspezifische Strategien zu verstehen, wie beispielsweise die Nutzung der Konzeptwechselstrategie (engl. *Conceptual Change*: Demastes et al., 1996) zur Konfrontation zweier konkurrierender Ansichten. Themenspezifische Instruktionsstrategien sind explizit auf die Unterstützung der Schülerinnen und Schüler bei dem Verständnis expliziter wissenschaftlicher Konzepte (z. B. der Evolutionstheorie) ausgerichtet. Hier bieten sich Repräsentationen wie Illustrationen (z. B. Zeitstrahl),

THEORETISCHER HINTERGRUND

Modelle (z. B. homologe Organe), Beispiele (z. B. natürliche Selektion beim Birkenspanner) und Analogien (z. B. Umweltveränderungen und selektiver Druck) zur Unterstützung instruktionaler Strategien an (z. B. Magnusson et al., 1999; Shulman, 1986). Die Facette Wissen über das Naturwissenschaftsverständnis von Schülerinnen und Schülern beinhaltet das Wissen über die Lernbedingungen und das Wissen über Lernschwierigkeiten. Ersteres verweist auf die Voraussetzungen für ein erfolgreiches Lernen, wie zum Beispiel nötiges Vorwissen bei den Schülerinnen und Schülern, um dem Unterrichtsthema folgen sowie gestellte Aufgaben während des Unterrichts lösen zu können (Magnusson et al., 1999). Im Kontext der Evolutionstheorie müsste im Vorfeld genetisches Grundlagenwissen geschaffen werden, um beispielsweise den Prozess der natürlichen Selektion verstehen zu können. Unter dem Wissen über Lernschwierigkeiten verstehen Magnusson et al. (1999) das Wissen, inwiefern die Abstraktheit sowie Komplexität eines Lerngegenstands das Verständnis bei Schülerinnen und Schülern behindern können. Dazu gehört auch das Wissen über spezifische Fehlvorstellungen in einem Themenbereich. Fehlvorstellungen führen dazu, dass sich wissenschaftlich korrekte Denkweisen bei Schülerinnen und Schülern erst gar nicht entwickeln können. Innerhalb der Evolutionstheorie wurden zahlreiche Fehlvorstellungen identifiziert und beschrieben, welche einen Einfluss auf das Lehren und Lernen des Prozesses der natürlichen Selektion haben können (detaillierte Erläuterungen s. Kapitel 2.2.2). Erst wenn die Lehrkräfte spezifische Fehlvorstellungen erkennen, ist die Ausgangslage geschaffen, mithilfe von Instruktionsstrategien ein erfolgreiches Lernen zu initiieren.

Die genannten Kernfacetten des FDW werden auch in weiteren Arbeiten im Bereich der bildungswissenschaftlichen und naturwissenschaftsdidaktischen Forschung thematisiert (z. B. Bravo und Cofré, 2016; Großschedl et al., 2014, 2015a, 2015b; Jüttner et al., 2013; Lee und Luft, 2008; Mahler et al., 2017; Paulick et al., 2016; Riese und Reinhold, 2009; Schmelzing et al., 2013) bzw. in Projekten behandelt (z. B. *COACTIV*: Kunter et al., 2011; *MT21*: Blömeke et al., 2008; *ProwiN*: Tepner et al., 2012). Ein aktuelles Modell zur Konzeptualisierung des Professionswissens mit einem starken Fokus auf das FDW bietet das *Refined Consensus Modell of PCK* (RCM: Carlson und Daehler, 2019), welches auf der Zusammenarbeit internationaler Expertinnen und Experten der fachdidaktischen Forschung basiert. Das RCM greift Aspekte des zuvor entwickelten *Consensus Model of PCK* (CM: Gess-Newsome, 2015) auf und führt neben der professionellen Wissensbasis von Lehrerinnen und Lehrern (FW, PW, Wissen der Schülerinnen und Schüler, Beurteilungswissen und Lehrplanwissen), das themenspezifische professionelle Wissen (Wissen über: Instruktionsstrategien, Repräsentationen, Schülervorstellungen, naturwissenschaftliche Praxis und Einstellungen) ein. Aus der professionellen Wissensbasis von Lehrerinnen und Lehrern und dem

themenspezifischen professionellen Wissen sowie unter Mitwirken verschiedener Faktoren wie dem Lernkontext und unterschiedlichen Beteiligten (Lehrerinnen und Lehrer, Schülerinnen und Schüler, Wissenschaftlerinnen und Wissenschaftler) formt sich das FDW. Dieses kann weiter in kollektives FDW (engl. *collective PCK: cPCK*), persönliches FDW (engl. *personal PCK: pPCK*) und ausgeführtes FDW (engl. *enacted PCK: ePCK*) differenziert werden (detaillierte Erläuterungen s. Kapitel 2.1.2.). Eine Besonderheit des RCM (Carlson und Daehler, 2019) ist im Gegensatz zu den Konzeptualisierungen von Magnusson et al. (1999) oder Baumert und Kunter (2006), welche sich nur auf die Nennung und Strukturierung von Wissenskomponenten bzw. Kompetenzen konzentrieren, der erweiterte Fokus auf der Wechselwirkung von unterrichtsrelevanten Faktoren und der Differenzierung zwischen theorie- und praxisrelevantem Wissen.

Wie von Shulman (1986) bereits definiert, beinhaltet FDW sowohl professionelles Wissen aus den Bereichen FW als auch PW. Hierbei gibt es unterschiedliche Annahmen darüber, inwiefern diese drei Wissensbereiche miteinander interagieren. Prinzipiell werden zwei Modelle unterschieden: das integrative und das transformative Modell (Gess-Newsome, 1999). Das integrative Modell geht davon aus, dass FDW als eigenständiger Wissensbereich nicht existiert, sondern eine Schnittmenge von FW und PW darstellt. Diese Annahme beinhaltet, dass sich FDW auch nicht empirisch trennen lässt (Gess-Newsome, 1999). Beim transformativen Modell wird das FDW als einzigartiger bzw. eigenständiger Wissensbereich zusammengefasst und dabei als Synthese aus FW und PW verstanden (Gess-Newsome, 1999). Dementsprechend ist FDW themenspezifisch und muss auf die jeweilige Zielgruppe ausgerichtet sein (Gess-Newsome, 1999). Ein Literaturüberblick von Chan und Hume (2019) über Arbeiten zum professionellen FDW zeigt, dass am häufigsten das transformative Modell zur theoretischen Rahmung für Studien genutzt wird.

Im vorliegenden Dissertationsprojekt wurde sich der Unterteilung des Professionswissens in FW, FDW und PW angeschlossen. Wie im *COACTIV*-Modell (Baumert und Kunter, 2006) konzeptualisiert, differenzieren sich die drei Wissensbereiche weiter in spezifische Wissensfacetten. Hierbei liegt der Fokus in den durchgeführten Studien auf dem Wissen über Evolutionsmechanismen (FW), dem Wissen über Schülerfehlvorstellungen (FDW) und dem pädagogischen Beurteilungswissen (PW). Besonders berücksichtigt wird das Wissen über Schülerfehlvorstellungen, da es übereinstimmend als eine Kernfacette des FDW identifiziert wurde (z. B. Park und Oliver, 2008). Damit ist es maßgeblich an diagnostischen Aktivitäten der Lehrerinnen und Lehrer hinsichtlich erbrachter Schülerleistungen im Biologieunterricht beteiligt ist (z. B. Förtsch et al., 2018).

2.1.2 Struktur und Konzeptualisierung von Professionswissen in Wissenstypen

In der kognitions- sowie pädagogisch-psychologischen Forschung sorgt die Strukturierung von Wissen immer wieder für Diskussionen darüber, inwiefern diese kognitiven Strukturen im menschlichen Gedächtnis gespeichert und abgerufen werden können (z. B. Anderson, 1986; Baddeley, 1992; Fenstermacher, 1994; Terhart, 1991). Hingegen herrscht Konsens darüber, dass das Wissen nicht als eindimensionales Konstrukt zu fassen ist, sondern aus mehreren Subsystemen besteht – den sogenannten Wissensarten bzw. -typen¹. In den zurückliegenden Jahrzehnten wurden hierzu zahlreiche Konzeptualisierungen und Begriffe eingeführt, um unterschiedliche Wissenstypen zu identifizieren und zu definieren (z. B. Baumert und Kunter, 2011a; Fenstermacher, 1994; Gramzow et al., 2013; Neuweg, 2011; Terhart, 1991; Tepner et al., 2012). Beispielsweise spricht Fenstermacher (1994) in seinen Ausführungen vom theoretisch-formalen und praktischen Wissen. Im Folgenden wird diese Differenzierung in weiteren Forschungsvorhaben verfolgt und zur theoretischen Fundierung genutzt (z. B. Baumert und Kunter, 2011a; Thiel, 2007). Eine andere grundlegende Konzeptualisierung von Wissenstypen stellt die Unterscheidung von deklarativem versus prozeduralem versus konditionalem bzw. strategisch-schematischem Wissen dar (z. B. Alexander and Judy, 1988; Anderson, 1976; Krathwohl, 2002; Paris et al., 1983; Shavelson et al., 2005; Tepner et al., 2012). Auch Shulman (1986) hat in seiner initialen Beschreibung des Professionswissens zwischen verschiedenen Wissenstypen – dem propositionalen Wissen (Faktenwissen), dem prozeduralen Wissen (Handlungswissen) und dem strategischen Wissen (Handlungswissen) – differenziert. Weitere Einordnungen unterschiedlicher Wissenstypen werden beispielhaft aufgeführt:

- Träges Wissen (Whitehead, 1929)
- Explizites vs. implizites Wissen (Polanyi, 1966)
- Semantisches und episodisches Wissen (Tulving, 1972)
- Prozedurales Interpretationswissen (Reif, 1987)
- Situationsbedingtes Wissen (De Jong & Ferguson-Hessler, 1996)
- Prozedurales, faktisches, konzeptuelles und metakognitives Wissen (Krathwohl, 2002)

Trotz zahlreicher Ordnungsversuche von Wissenstypen besteht in der aktuellen Forschungslandschaft überwiegend Einigkeit darüber, inwiefern Wissen strukturiert werden kann, wobei sich die Differenzierung in deklaratives und prozedurales Wissen etabliert hat (z. B. Anderson und Lebière, 1998; Bromme, 2001; De Jong und Ferguson-Hessler, 1996; Edelmann, 2000; König et al., 2014; Lukesch, 2001; Rost et al., 2004; Schnotz, 1994; Seel, 2003; Steiner, 2001).

¹ In den weiteren Ausführungen wird ausschließlich von Wissenstypen gesprochen.

Deklaratives Wissen. Als deklaratives Wissen wird das Wissen über Fakten, Begriffe und Prinzipien, aber auch deren Vernetzung zu konzeptuellem Wissen verstanden, welches einem Individuum ermöglicht, die Umwelt wahrzunehmen und zu interpretieren (z. B. Krathwohl, 2002; Baumert und Kunter, 2011a; Schnotz, 1994; Tepner et al., 2012). Hierbei ist dieses Wissen dem Bewusstsein potenziell zugänglich – auch explizites Wissen genannt – und kann unmittelbar beschrieben werden (Polanyi, 1966; Schnotz, 1994). Maßgeblich für das Erlernen und Abrufen deklarativen Wissens ist die Auseinandersetzung mit Verstehensprozessen, beispielsweise während des akademischen Diskurses (Baumert und Kunter, 2011a; Schnotz, 1994). Shavelson et al. (2005) beschreiben die unterschiedlichen Wissenstypen (deklaratives, prozedurales und konditionales Wissen) als geordnete Menge, welche mit zunehmender Stufe dem Grad der Fähigkeit einer Person entsprechen. Hier ist das deklarative Wissen (Wissen, dass) der ersten Stufe der Fähigkeitsskala zugeordnet und markiert damit den Startpunkt der Professionalisierung.

Prozedurales Wissen. Das prozedurale Wissen wird grundlegend als Handlungswissen verstanden, wobei die subsumierten Ausprägungen wie zum Beispiel das Wissen über Strategien oder Methoden und Prozeduren sehr heterogen sein können (z. B. Lukesch, 2001; Jüttner und Neuhaus, 2013; Tepner et al., 2012). Hierbei handelt es sich um handlungsrelevantes Wissen, also das „Wissen wie“, welches notwendig ist, um komplexe kognitive Prozesse sowie motorische Handlungen auszuüben (Seel, 2003). Da die entsprechenden Handlungsabläufe zum Teil unbewusst bzw. intuitiv erfolgen, ist eine sprachliche Formulierung des prozeduralen Wissens häufig schwierig (Baumert und Kunter, 2013). Dementsprechend wird prozedurales Wissen auch als implizites Wissen bezeichnet (Polanyi, 1966; Seel, 2003). Prozedurales Wissen umfasst neben automatisierten Handlungsabfolgen auch Prozesse, welche einem Schritt-für-Schritt Schema folgen (De Jong und Ferguson-Hessler, 1996). In diesem Fall ist die Handlung nicht zwangsläufig unterbewusst bzw. implizit, sondern basiert auf begründeten Entscheidungen des Individuums, um ein bestimmtes Ziel zu erreichen (Anderson, 2001).

Einerseits zeigt die Konzeptualisierung von deklarativem und prozeduralem Wissen, dass beide Wissenstypen voneinander abgegrenzt werden können und andererseits, dass beide Wissenstypen sich gegenseitig bedingen. Das deklarative Wissen ist der Ausgangspunkt bzw. die Voraussetzung für prozedurales Wissen (z. B. Anderson, 1983; Shavelson et al., 2005). Demnach muss vorerst deklaratives Wissen erworben werden, welches später in konkreten Handlungssituationen, also über regelmäßige Übung und Kontextualisierung, in prozedurales Wissen transformiert werden kann (z. B. Anderson, 1982, 2001; Blömeke et al., 2010; Schneider und Stern, 2010).

Die entwickelten kognitions- und pädagogisch-psychologischen Ansätze zur Wissensstrukturierung in deklaratives und prozedurales Wissen wurden in der mathematik- und naturwissenschaftlichen Forschung übernommen bzw. leicht modifiziert (z. B. Jüttner

und Neuhaus, 2013; Olszewski et al., 2010; Star und Stylianides, 2013; Tepner et al., 2012). Hierzu wurden zahlreiche Forschungsprojekte initiiert, welche den Anspruch hatten, das professionelle Wissen (angehender) Lehrkräfte adäquat zu modellieren (z. B. Borowski et al., 2010, 2011; Kröger et al., 2013; Lange et al., 2012; Olszewski, 2010a; Riese und Reinhold, 2010; Tepner et al., 2012). Im Folgenden werden Studien aus der naturwissenschafts- und mathematikdidaktischen Forschung vorgestellt, welche neben den bereits thematisierten Wissensbereichen auch eine Differenzierung zwischen verschiedenen Wissenstypen vornehmen.

In dem Forschungsprojekt *ProwiN (Professionswissen in den Naturwissenschaften)* wird das Professionswissen von Lehrkräften der Naturwissenschaften untersucht (Kirschner et al., 2017). Hierfür wird ein Modell zur Entwicklung von Items für einen Papier-Bleistift-Fragebogen vorgestellt, das neben den Wissensbereichen des professionellen Wissens (FW, FDW und PW) auch die Wissenstypen (deklarativ und prozedural) in den drei Naturwissenschaften (Biologie, Chemie und Physik) abbildet. Deklaratives Wissen wird innerhalb von *ProwiN* als das Wissen über Begriffe, Fakten und Prinzipien operationalisiert. Das Wissen über Praktiken und Prozesse im Unterricht wird dem prozeduralen Wissen zugeordnet (Fischer et al., 2012). Weiterhin wird innerhalb des Modells das konditionale Wissen, das nach Anderson (1982) und Paris et al. (1983) eine Erweiterung des prozeduralen Wissens darstellt, als weiterer Wissenstyp ergänzt. Das Modell zur Entwicklung von Test-Items aus *ProwiN* sowie dem daraus resultierenden Fragebogen der Vorstudie adaptieren Borowski et al. (2011), um eine Vergleichsstudie zum Fachwissen von Studierenden, Referendaren und Lehrenden im Fach Physik durchzuführen. Innerhalb der Studien wird kein expliziter Bezug der Ergebnisse auf die zuvor thematisierte Differenzierung zwischen deklarativem und prozeduralem Wissen genommen.

In der Studie von Schmelzing et al. (2010) wird ein Messinstrument (videogestützter Papier-Bleistift-Fragebogen) entwickelt, welches das FDW von angehenden und praktizierenden Biologielehrkräften erfassen soll. Abweichend zur Differenzierung in deklaratives und prozedurales Wissen aus der kognitionspsychologischen Forschung führen Schmelzing et al. (2010) neue Begriffe ein – „FDW über Aktion“ (engl. *PCK-on-action*), „FDW in Aktion“ (engl. *PCK-in-action*) und „Reflektion“ (engl. *reflection*). Hierbei bezieht sich das *PCK-on-action* auf das theoretisch-formale, explizite und deklarative Wissen. *PCK-in-action* kann als das implizite, prozedurale Wissen und die Fähigkeiten beschrieben werden. Um beide Wissenstypen zu integrieren, sind reflektive Komponenten (*reflection-on-action* und *reflection-in-action*) notwendig. Innerhalb der Studie fokussieren Schmelzing et al. (2010) lediglich das *PCK-on-action* und *reflection*, wohingegen prozedurale Facetten (*PCK-in-action*) des Wissens nicht weiter untersucht werden.

Analog zu der Studie von Schmelzing et al. (2010) entwickelten Olszewski et al. (2010) ein weiteres Erhebungsinstrument, um das deklarative und prozedurale FDW bei Physiklehrkräften zu untersuchen. Für die beiden Wissenstypen werden zwei unterschiedliche Messverfahren verwendet. Zum einen ein Fragebogen, mit welchem vor allem deklaratives Wissen abgedeckt wird und zum anderen Videoanalysen, welche zur Erfassung des umgesetzten prozeduralen Wissens dienen. Innerhalb der Studie wird das deklarative FDW als Wissen über Fehlvorstellungen, Wissen über den Lehrplan und Abschätzung von Schwierigkeiten operationalisiert. Zu dem prozeduralen FDW gehörte das gezeigte Wissen zur geeigneten Auswahl von Aufgaben, Aufgabenzuweisung während der Unterrichtsstunde, Reaktion auf die Antworten der Schülerinnen und Schüler und Vergabe von Hausaufgaben. Die Kombination beider Messinstrumente soll Aufschluss darüber geben, inwiefern theoretisches Wissen angehender Lehrkräfte in praktischen Unterrichtssituationen implementiert wird.

Im Projekt *KiL (Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen)* wird ein Modell etabliert, das als Grundlage für die Entwicklung der Fragebogenitems dient. Innerhalb des biologischen Teilprojekts wird zwischen den kognitiven Prozessen „erinnern und abrufen“ (engl. *remember and retrieve*) und „anwenden und verstehen“ (engl. *apply and understand*) unterschieden (Großschedl et al., 2015a), während in der Physikdidaktik die drei Wissenstypen: deklaratives, prozedurales und schematisch-strategisches Wissen verwendet werden. In Anlehnung an Tamir (1988), welcher zwischen Wissen und Fertigkeiten differenziert, wird *remember and retrieve* als deklaratives Wissen verstanden, wohingegen *apply and understand* als prozedurales Wissen operationalisiert wird. Zentral ist die kognitionspsychologische Annahme, dass Wissen aus dem Langzeitgedächtnis in das Arbeitsgedächtnis überführt wird und sich in einer konkreten Handlungssituation artikuliert (Krathwohl, 2002). Eine ähnliche Konzeptualisierung von Wissen wird ebenfalls von Blömeke et al. (2010) sowie König et al. (2011) in *TEDS-M* (engl. *Teacher education and development study in mathematics*) eingeführt.

Förtsch et al. (2018) empfehlen zur Strukturierung von professionellem Wissen ein zwei-dimensionales Modell, in welchem neben den Wissensbereichen FW, FDW und PW weiterhin in drei Wissenstypen unterschieden wird: dem „Wissen, dass“ (engl. *KT: knowing that*), welches das Wissen zu Fakten, Konzepten und Prinzipien zusammenfasst, dem „Wissen, wie“ (engl. *KH: knowing how*), das Wissen über Aktionen, Prozeduren und Manipulationen, sowie dem „Wissen, wann und warum“ (engl. *KW: knowing when and why*), welches sich auf die Anwendung beider Wissenstypen bezieht. Hierbei umfassen das *KH* und *KW* das handlungsbezogene Wissen, also das Wissen, welches in Handlungen besonders relevant wird. Ziel dieser Konzeptualisierung war es, die Anschlussfähigkeit an die kognitionspsychologische sowie die bildungs- und naturwissenschaftliche Forschung

THEORETISCHER HINTERGRUND

herzustellen, und somit Studienergebnisse vergleichbar zu machen. Weiterhin wurde explizit thematisiert, welche Facetten des Professionswissens innerhalb der Fächer Mathematik und Biologie den entsprechenden Wissenstypen zuzuordnen sind. Beispielsweise ist das Wissen über Schülerfehlvorstellungen im Biologieunterricht dem *KT* in der Wissensfacette FDW zugeordnet, wohingegen bei der Auswahl einer geeigneten Instruktionsstrategie zur Überwindung der Schülerfehlvorstellung das *KH* im Wissensbereich FDW benötigt wird.

Eine ganz andere Herangehensweise an die Konzeptualisierung von handlungsnahem Professionswissen eröffnet das *Refined Consensus Model of PCK (RCM)*, welches aus einem Zusammentreffen US-amerikanischer Wissenschaftlerinnen und Wissenschaftler der Fachdidaktik resultierte (Carlson und Daehler, 2019; s. Abbildung 2.3). Das *RCM* bietet Forschenden eine differenzierte Analyse des FDW von Lehrkräften, die über deklaratives Wissen hinausgeht und sich auf die dynamische Anwendung in handlungsorientierten Situationen konzentriert. Im Kern unterscheidet das Modell drei verschiedene Bereiche des FDW - das *cPCK*, *pPCK* und das *ePCK*. Das *cPCK* stellt den Konsens des FDW mehrerer Personen (z. B. Pädagoginnen und Pädagogen, Lehrerinnen und Lehrern) in einem bestimmten Fachbereich dar. Dies führt zu einer spezialisierten Wissensbasis für den naturwissenschaftlichen Unterricht, die von Fachleuten verwendet wird und sich auf spezifische Schülerinnen und Schüler in einem spezifischen Lernkontext bezieht. Somit ist das *cPCK* nicht privat, sondern eher öffentlich und umfasst das Wissen mehrerer Personen. Die Schicht zwischen dem *cPCK* und dem *pPCK* wird als Lernkontext bezeichnet. Das Lehren und Lernen von Naturwissenschaften findet immer in einem spezifischen Kontext statt, der von zahlreichen Faktoren beeinflusst wird, wie zum Beispiel der Bildungspolitik (z. B. Bundespolitik, nationale Standards), den individuellen Eigenschaften der Schüler (z. B. Alter, Sprachkenntnisse) und der spezifischen Lernumgebung (z. B. Klassenzimmer, Lernen außerhalb der Schule). Der Lernkontext ist der Vermittler, der die Handlungen und Entscheidungen der Lehrkräfte beeinflusst. Das *pPCK* ist die Quelle von Wissen und Fertigkeiten, die eine Lehrkraft während der Unterrichtspraxis nutzen kann. Diese Wissensbasis wird zum einen durch das *cPCK* gespeist, zum anderen aber auch durch Erfahrungen im Unterricht und Beiträge von Peers. Das *ePCK* ist eine Teilmenge des *pPCK*, die das spezifische Wissen und die Fertigkeiten von Lehrkräften in bestimmten Unterrichtssituationen darstellt. Das *ePCK* umfasst die Planung von Unterricht (engl. *plan*), das Unterrichten selbst (engl. *teach*) und die Reflexion des Unterrichts (engl. *reflect*). Dabei artikuliert sich das *ePCK* vor allem in der Interaktion zwischen Lehrer und Schülern im Unterricht (z. B. Wahl der Unterrichtsstrategien, diagnostisches Wissen), aber auch in den Schülerergebnissen der Interaktion (Beurteilungswissen).

Weiterhin berücksichtigt das Modell ebenfalls die Bereiche des professionellen Wissens von Lehrkräften (d.h. FW, FDW, PW, Lehrplanwissen, Beurteilungswissen und Wissen der Schüler).

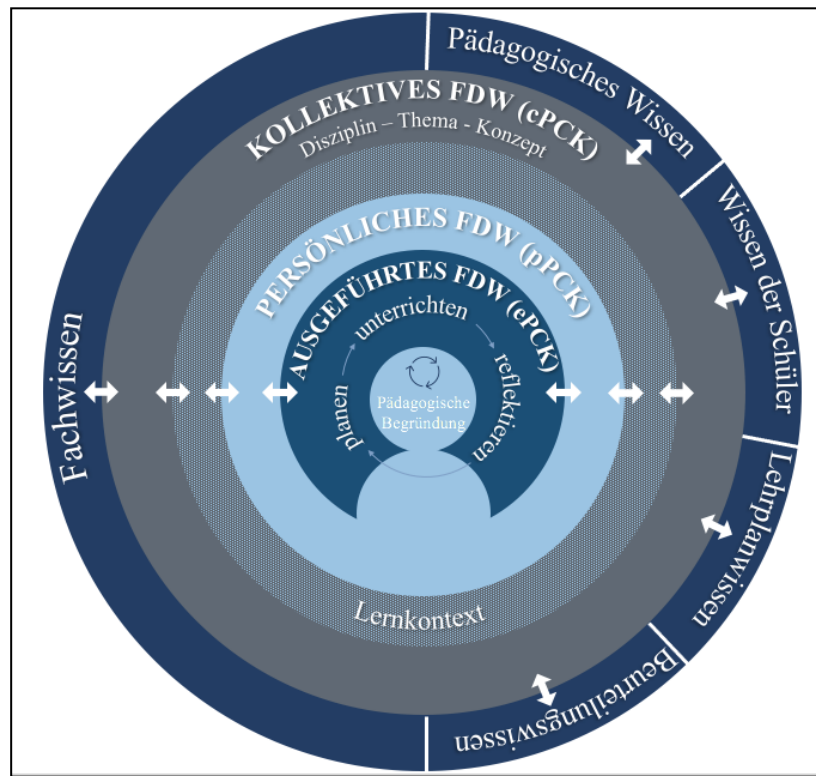


Abbildung 2.3 Refindet Consensus Modell of PCK (verändert nach Carlson und Daehler, 2019).

In den durchgeführten Studien der vorliegenden Dissertation wurde sich der grundlegenden Konzeptualisierung der Wissenstypen in deklaratives und prozedurales Wissen angeschlossen, da hier ein Konsens in der kognitionspsychologischen Forschung herrscht (z. B. De Jong und Ferguson-Hessler, 1996). Auch in bildungswissenschaftlichen und naturwissenschaftsdidaktischen Studien wurde diese Konzeptualisierung verwendet, wobei terminologische Unterschiede zu beobachten sind (z. B. Schmelzing et al., 2010). In einer Studie wurde explizit das handlungsorientierte FDW fokussiert, wodurch sich hier das relativ neue *RCM* (Carlson und Daehler, 2019) als geeignete theoretische Rahmung angeboten hat. Besonders die deutliche Differenzierung des FDW in unterschiedliche Bereiche (*cPCK*, *pPCK*, *ePCK*), welche sich hinsichtlich ihrer Handlungsorientierung unterscheiden, markieren eine zentrale Neuerung in Abgrenzung zu anderen bisher etablierten Modellen zum FDW (z. B. Magnusson et al., 1999). Zusätzlich motivierend war die Möglichkeit, dieses kürzlich entwickelte Modell zur Konzeptualisierung des FDW in einer empirischen Studie einzusetzen.

2.2 Die biologische Evolution und Evolutionstheorie

Die Evolution beschreibt den Prozess der Veränderungen in Populationen oder Taxa von Organismen im Laufe der Zeit. Weiterhin liefert sie eine Erklärung für die biologische Vielfalt unserer Umwelt und die Ähnlichkeiten zwischen Organismen. Wissenschaftlerinnen und Wissenschaftler sind sich einig, dass die natürliche Selektion, der Genfluss und der genetische Drift die primären Mechanismen der Evolution sind (z. B. Lynch, 2007; Mayr, 1982). Ungeachtet der Komplexität der Evolutionstheorie lassen sich die Kernideen in wenigen Sätzen zusammenfassen. Die biologische Evolution kann auf der untersten Ebene (Mikroevolution) als Veränderungen in der Allelfrequenz einer Population definiert werden. Diese Veränderungen und die daraus resultierende genetische Variation werden durch Rekombination und Mutation hervorgerufen. Die Variabilität ist die Grundlage für die natürliche Selektion, welche den Anpassungsprozess innerhalb der Arten beschreibt. Diejenigen Individuen, die genetisch besser an die Umweltbedingungen angepasst sind, haben eine höhere Überlebenswahrscheinlichkeit und dementsprechend auch eine gesteigerte Reproduktionsrate. In der nächsten Generation treten diese vorteilhaften Merkmale häufiger auf, wodurch die Population Angepasstheiten an die neuen Umweltbedingungen zeigen (z. B. Andersson und Wallin, 2006).

In der vorliegenden Dissertation wird der Prozess der natürlichen Selektion fokussiert, da dieser häufig zu Verständnisproblemen der biologischen Evolution bei Schülerinnen und Schülern führt. Dementsprechend werden vorerst die Probleme beim Lehren und Lernen der biologischen Evolution beschrieben (siehe Kapitel 2.2.1), um anschließend spezifische Fehlvorstellungen zur natürlichen Selektion zu thematisieren (siehe Kapitel 2.2.2).

2.2.1 Lehren und Lernen der Evolution

Für das Lehren und Lernen der Biologie ist die Evolutionstheorie ein essenzieller Bestandteil. Sie ist das übergeordnete Organisationsprinzip der Biologie und verbindet einzelne Fakten zu einem umfassenden Erklärungsrahmen. Hierbei sind evolutionäre Prozesse, wie beispielsweise die natürliche Selektion, die Grundlage aller biologischen Themen und ermöglicht ein konzeptuelles Verständnis (z. B. Basel et al., 2014; Bishop und Anderson, 1990; Furtak, 2012; Opfer et al., 2012). Erst dieses konzeptuelle Verständnis der Evolutionstheorie ermöglicht die Teilnahme am gesellschaftlichen Diskurs aller Beteiligten, beispielsweise bei der Risikobeurteilung von Virusinfektionen und deren Potenzial, durch Selektionsmechanismen neue Varianten hervorzubringen (Nationale Akademie der Wissenschaften Leopoldina, 2017). Folglich ist die Evolutionstheorie ein zentraler Bestandteil der Lehrpläne des naturwissenschaftlichen Unterrichts in zahlreichen Ländern (z. B. Deutschland: Ständige Konferenz der Kultusminister der Länder in der

Bundesrepublik Deutschland [KMK] 2005; USA: Next Generation Science Standards [NGSS] 2013). Trotz dieser Fokussierung der Evolutionstheorie in den Schulen beschreiben zahlreiche Studien große Probleme beim Lehren und Lernen der Evolution (z. B. Boujaoude et al., 2011; Gregory, 2009; Großschedl et al., 2014; Lucero et al., 2019; Nehm und Reilly, 2007; Nehm und Schonfeld, 2007, 2008). In diesem Kontext werden zahlreiche wissenschaftlich inadäquate Denkweisen als bedeutende Lernhindernisse beschrieben. Bekannt ist, dass Schülerinnen und Schüler mit stark verankerten fehlerbehafteten Vorstellungen in den naturwissenschaftlichen Unterricht kommen, welche genutzt werden, um evolutionäre Prozesse und Mechanismen vorherzusagen und zu erklären (Coley und Tanner, 2012; Sinatra et al., 2008). Obwohl diese Erklärungen im Alltag der Schülerinnen und Schüler gut funktionieren, unterscheiden sie sich von den wissenschaftlich akzeptierten Erklärungen und werden daher als fehlerhaft angesehen (Yip, 1998). Diese Vorstellungen werden beispielsweise als Fehlvorstellungen, Alltagsvorstellungen, alternative Vorstellungen oder naive Ideen bezeichnet (z. B. Graf und Hamdorf, 2011; Leonard et al., 2014; Maskiewicz und Lineback, 2013). Abgesehen von einigen Unterschieden in der kontextabhängigen Begriffsverwendung basieren sie jedoch alle auf der Grundannahme ungenauer Vorstellungen oder Ideen zu einem naturwissenschaftlichen Phänomen². Besonders problematisch ist die Tatsache, dass wissenschaftlich korrekte Denkweisen mit Fehlvorstellungen koexistieren können (z. B. Fosnot und Perry, 2005; Palmer, 1999; Shtulman & Valcarcel, 2012). Dementsprechend kann auch nicht das Erlernen eines wissenschaftlichen Konzepts per se wissenschaftlich inadäquate Denkweisen beseitigen.

Fehlvorstellungen von Schülerinnen und Schülern resultieren beispielsweise aus den eigenen Erfahrungen, der wissenschaftlich ungenauen Sprache von Lehrerinnen und Lehrern im Unterricht oder verwendeten fehlerhaften Schul- bzw. Sachbüchern (z. B. King, 2010; Nehm et al., 2009). Unabhängig davon können Fehlvorstellungen über Naturphänomene vor der Schulzeit und somit bereits im jungen Alter entstehen, wenn Kinder ihre physische und soziale Umwelt erkunden (z. B. Beardsley, 2004; Bruckermann et al., 2021; Byrne et al., 2009; Driver, 1988; Evans, 2000). Folglich werden die im frühen Kindesalter entwickelten Fehlvorstellungen auch in Zukunft genutzt, um evolutionäre Prozesse zu erklären. So zeigt die Forschung, dass Fehlvorstellungen zur Evolution auch bei Schülerinnen und Schülern (z. B. Baalman et al., 2004; Demastes et al., 1995; Zabel und Gropengießer, 2011), Biologiestudierenden (z. B. Beggrow und Nehm, 2012; Bishop und Andersen, 1990; Dagher und BouJaoude, 1997; Graf und Soran, 2011; Nehm und Reilly, 2007; Nehm und Schonfeld, 2008; Weitzel und Gropengießer, 2009; Yates und

² In weiteren Ausführungen wird nach Graf und Hamdorf (2011) ausschließlich von Fehlvorstellungen gesprochen.

Marek, 2014), Medizinstudierenden (Brumby, 1984) und Lehrerinnen und Lehrern der Naturwissenschaften (Nehm und Schonfeld, 2007) existieren.

Die Tatsache, dass die Prinzipien der Evolution sowohl von Kindern, Schülerinnen und Schülern und Studierenden als auch von großen Teilen der Gesellschaft weitestgehend missverstanden werden, hat Wissenschaftlerinnen und Wissenschaftler dazu motiviert, Fehlvorstellungen zur Evolution zu identifizieren und Unterrichtsstrategien zu finden, um diesen adäquat begegnen zu können (z. B. Gregory, 2009; Harms und Reiss, 2019; Ziadie und Andrews, 2018).

2.2.2 Fehlvorstellungen zur natürlichen Selektion

Die natürliche Selektion ist der zentrale Prozess des evolutionären Wandels, welcher zu adaptiven Merkmalsveränderungen bei Organismen führt. Aus der Forschung ist bekannt, dass der Prozess der natürlichen Selektion eines der am schwierigsten zu vermittelnden Themen in der Biologie ist (z. B. Bishop und Anderson, 1990; Nehm und Reilly, 2007). Hierbei werden zahlreiche spezifische Fehlvorstellungskategorien beschrieben, welche das Erlernen des Prozesses der natürlichen Selektion erschweren bzw. verhindern können (Gregory, 2009). Diese spezifischen Fehlvorstellungen sind äußerst vielfältig und können unterschiedliche Bereiche des Prozesses der natürlichen Selektion betreffen. Am häufigsten basieren Fehlvorstellungen auf der Annahme, dass Evolution stattfindet, weil sich einzelne Organismen während ihres Lebens aktiv weiterentwickeln und an neue Umweltbedingungen willentlich anpassen (z. B. Bishop und Anderson, 1990; Settlage, 1994; Nehm und Schonfeld, 2008). In diesem Zusammenhang zeigen Forschungsarbeiten der vergangenen Jahre, dass vor allem anthropomorphe, teleologische und lamarckistische Fehlvorstellungen genutzt werden, um den Prozess der natürlichen Selektion zu erklären (z. B. Bishop und Anderson, 1990; Bizzo, 1994; Demastes et al., 1995; Ferrari und Chi, 1998; Hammann und Nehm, 2020; Jensen und Finley, 1996; Jiménez-Alexandre et al., 1992; Jungwirth, 1977; Kampourakis und Zogza, 2008; Moore et al., 2002; Nehm et al., 2009; Nehm und Reilly, 2007; Nehm und Schonfeld, 2007; Prinou et al., 2008; Settlage, 1994; Tamir und Zohar, 1991). Diese prominenten Fehlvorstellungskategorien sind Grundlage für die durchgeführten Studien dieser Dissertation und werden im Folgenden kurz erläutert:

Anthropomorphe Fehlvorstellung: Anthropomorphe Erklärungen sind dadurch gekennzeichnet, dass menschliche Eigenschaften, wie zum Beispiel das Denken oder Fühlen, auf nichtmenschliche Organismen übertragen werden (z. B. Byrne et al., 2009; Southerland et al., 2001). Hierbei wird Tieren, Pflanzen oder Bakterien ein Bewusstsein zugeschrieben, welches ihnen erlaubt, die eigene Unangepasstheit wahrzunehmen. Die daraus initiierte Veränderung eines Merkmals ist dementsprechend das Ergebnis von absichtlichen und zielgerichteten Verhalten oder Handlungen, welche aktiv von den

Individuen gesteuert werden, um sich den veränderten Umweltbedingungen anzupassen (z. B. Baalman et al., 2004; Demastes et al., 1995; Gregory, 2009; Kallery und Psillos, 2004; Tamir und Zohar, 1991). Auch die Natur oder die natürliche Selektion selbst können als bewusst handelnde Akteure betrachtet werden, welche die Entstehung von Angepasstheiten willentlich hervorrufen (Kampourakis und Zogza, 2008; Sinatra et al., 2008).

Lamarckistische Fehlvorstellung (Gebrauch oder Nichtgebrauch). Die lamarckistische Fehlvorstellung ist nach ihrem Namensgeber Jean-Baptiste Lamarck benannt und basiert auf der von ihm entwickelten Evolutionstheorie. Diese Fehlvorstellung bezieht sich auf die Tendenz von Schülerinnen und Schülern, den Prozess der natürlichen Selektion durch den Gebrauch bzw. Nichtgebrauch von Organen oder Fähigkeiten zu beschreiben. Die daraus resultierende Merkmalsveränderung wird entsprechend Lamarck an die nächste Generation weitervererbt (z. B. Andrews et al., 2011; Gregory, 2009; Kalinowski et al., 2016; Kampourakis und Zogza, 2008; Nehm et al., 2009; Prinou et al., 2008).

Teleologische Fehlvorstellung: Bei teleologischen Erklärungen wird die Entwicklung eines Merkmals als zielgerichtet und zweckmäßig angesehen. Merkmale entstehen oder gehen verloren, weil diese zu einem Überlebensvorteil des Individuums führen (z. B. Alters und Nelson, 2002; Andrews et al., 2011; Beardsley, 2004; Bishop und Anderson, 1990; Kampourakis et al., 2012; Kampourakis und Zogza, 2008; Nehm et al., 2009; Nehm und Reilly, 2007; Nehm und Schonfeld, 2007, 2008; Settlage, 1994; Sinatra et al., 2008; Southerland et al., 2001). Hierbei folgen teleologische Fehlvorstellungen immer einem „Start-Ziel-Schema“, wonach der Zweck eine Merkmalsveränderung initiiert und zu einem unveränderlichen Endergebnis führt (Stover und Mabry, 2007). Die natürliche Selektion wird nach der teleologischen Fehlvorstellung nicht als Prozess, sondern eher als singuläres Ereignis verstanden (Sinatra et al., 2008).

Ergebnisorientierte Erklärungen der natürlichen Selektion sind inhärente Tendenzen des Menschen und basieren intuitiv auf den eigenen persönlichen Erfahrungen im ziel- und problemorientierten Denken (z. B. Gregory, 2009). Diese Tendenzen werden dadurch verstärkt, dass Schülerinnen und Schüler im naturwissenschaftlichen Unterricht häufig dazu aufgefordert werden, Naturphänomene kausal zu erklären (Olander, 2009). Zusätzlich werden diese Denkweisen durch den alltäglichen, aber auch wissenschaftlichen Sprachgebrauch unterstützt (Alters und Nelson, 2002; Nehm et al., 2010). Wird beispielsweise von der „Anpassung eines Organismus“ gesprochen, ohne diese Aussage in einen expliziten Kontext einzubetten, kann eine Zweck- bzw. Zielorientiertheit des Organismus suggeriert werden (z. B. Baalman et al., 2004; Gregory, 2009). Das Ziel des Biologieunterrichts sollte es daher sein, Schülerinnen und Schüler dabei zu unterstützen, konzeptuelles Wissen zur Evolution zu erwerben (Gregory, 2009). Hierfür ist es wichtig,

dass Lehrerinnen und Lehrer die Fehlvorstellungen ihrer Schülerinnen und Schüler erkennen, um darauf aufbauend wissenschaftliche Denkweisen zu initiieren.

2.3 Simulierte Klassenraumumgebungen

Die Entwicklung von digitalen Instrumenten im Bereich des naturwissenschaftlichen Lernens und Lehrens ist in den vergangenen Jahren in den Fokus vieler Forschungsgruppen gerückt (z. B. Bradley und Kendall, 2014). Maßgeblich verantwortlich hierfür sind vor allem neue technologische Fortschritte und Möglichkeiten, wie zum Beispiel die flächendeckende Ausbreitung von Hochgeschwindigkeits-Internet sowie die gestiegene Verarbeitungsleistung moderner Computer. Aktuelle Forschungsprojekte zum naturwissenschaftlichen Unterricht sowie im pädagogisch-psychologischen Bereich reichen von der Entwicklung und Verwendung digitaler Fragebögen bis hin zu videobasierten Systemen und umfassenden digitalen Simulationen (z. B. *The Virtual Chemistry Laboratory*: Dalgarno et al., 2009; *EvoSketch*: Fiedler et al., 2018; *Interactive Lessons in Biology*: Heitz et al., 2010; *Das Schülerinventar*: Jansen et al., 2019, 2021; Kaiser et al., 2015). Viele dieser Projekte profitieren von neuen technologischen Innovationen, die es erlauben, große Datensätze über Lernprozesse und deren Ergebnisse gleichzeitig zu erfassen, zu verarbeiten und direkt in Statistik-Software zu überführen. Dadurch können zeitaufwändige Verfahren vermieden und personelle Ressourcen gespart werden. Gleichzeitig können digitale Instrumente wie beispielsweise simulierte Klassenraumumgebungen die Nutzerinnen und Nutzern mit einer realitätsnahen digitalen Lernumgebung konfrontieren, wodurch eine attraktive und interaktive wissenschaftliche Lernerfahrung geboten werden kann (z. B. Fiedler et al., 2002; Plass et al., 2009).

Im Folgenden werden simulierte Klassenraumumgebungen definiert und konzeptionelle Unterschiede beschrieben. Darauf aufbauend wird der Simulierte Klassenraum vorgestellt, welcher die Grundlage für die Studien dieser Dissertation darstellt (siehe Kapitel 2.3.1).

Neue digitale Innovationen bieten Wissenschaftlerinnen und Wissenschaftlern die technologischen Rahmenbedingungen, simulierte Klassenraumumgebungen zu entwerfen und diese als Werkzeug für die Forschung bereitzustellen (z. B. Bradley und Kendall, 2014; Cheong, 2010; Gurvitch und Metzler, 2009). Hierbei besitzen simulierte Klassenraumumgebungen das Potenzial, die Lücke zwischen realen Unterrichtssituationen in Klassenzimmern und der theoretischen Lehramtsausbildung an Universitäten zu schließen (Hixon und So, 2009). Als Simulationen werden vereinfachte dynamische Modelle der Realität bezeichnet, welche in digitale Systeme implementiert werden, aber dennoch eine allgemeine Gültigkeit besitzen (z. B. de Jong und van Joolingen, 1998; Sauvé

et al., 2007). Die Reduktion der Komplexität innerhalb von simulierten Klassenraumumgebungen ermöglicht zum einen, die kognitive Belastung der Lehrerinnen und Lehrer zu regulieren, und zum anderen, die Anwendung von professionellem Wissen unter kontrollierten Bedingungen zu beobachten (z. B. Cheong, 2010; Gurvitch und Metzler, 2009). Simulationen ermöglichen es den Nutzerinnen und Nutzern, ihr Wissen interaktiv anzuwenden und weiterzuentwickeln, indem sie Problemsituationen begegnen, Entscheidungen treffen und Handlungen ausprobieren, ohne dabei reale Konsequenzen wie zum Beispiel negatives Feedback durch die Schülerschaft zu erfahren (z. B. van Berkum und de Jong, 1991; Hixon und So, 2009; Plass et al., 2009). Weiterhin können innerhalb von Simulationen spezifische Szenarien beliebig häufig wiederholt werden, um diese über längere Zeiträume zu üben, als es in der Realität möglich wäre (Girod und Girod, 2008). Die meisten simulierten Klassenraumumgebungen basieren auf Rollenspielsimulationen, in welchen die Nutzerinnen und Nutzer in die Rolle einer am Unterricht beteiligten Person schlüpfen (Lehrkraft oder Schülerin bzw. Schüler) und dabei deren Aufgaben in diesem simulierten Szenario übernehmen (Veletsianos et al., 2010). Am häufigsten wird innerhalb der simulierten Klassenraumumgebung die Rolle der Lehrkraft übernommen, die in Interaktion mit den virtuellen Schülerinnen und Schülern tritt. Dabei werden reale, klassenspezifische Unterrichtssituationen simuliert, beispielsweise Unterrichtsstörungen durch „Hereinrufen“ oder „Herumhampeln“, auf welche die Lehrkraft adäquat reagieren muss (Clapper, 2010). Generell lassen sich simulierte Klassenraumumgebungen in immersive und nicht-immersive Simulationen unterscheiden. Bei Letzteren handelt es sich um Simulationen, welche die Charakteristika der virtuellen Schülerinnen und Schüler im Klassenraum durch Texte sowie Bilder darstellen. Bei immersiven Simulationen werden dynamisch-visuelle Darstellungsformen eines Klassenzimmers gewählt, in das Avatare integriert werden (Dalgarno et al., 2016). Die realistische Einbindung der Nutzerinnen und Nutzer durch einen dreidimensionalen Avatar ermöglicht eine stärkere Identifikation mit der zugewiesenen Rolle (Jamaludin et al., 2009). Im Gegensatz dazu haben nicht-immersive Klassenraumsimulationen den Vorteil, dass sie deutlich komplexreduzierte Umgebungen mit geringerer kognitiver Belastung bieten (Fiedler et al., 2007).

Bisher konzentrieren sich die meisten simulierten Klassenraumumgebungen in der Lehramtsausbildung auf die Facetten des PW, beispielsweise der effektiven Klassenführung und Leistungsbeurteilung, oder auf psychologische Fragestellungen nach der Selbstwirksamkeit oder interpersonalem Kompetenz (z. B. *TeachMETM*: Bautista und Boon, 2015; *simSchool*: Christensen et al., 2011; *VirtualPREX*: Dalgarno et al., 2016; *TeachLivE*: Dede, 2009; *ClassSim*: Ferry et al., 2004; *CS-TGCTS*: Yeh, 2004). Erste Befunde über den Einsatz simulierter Klassenraumumgebungen zeigen, dass die Lehramtsstudierenden kritischer über komplexe Unterrichtssituation nachdenken (Ferry et

al., 2004) oder schneller ein Gefühl der Selbstwirksamkeit im Vergleich zu traditionellen Ansätzen der Lehramtsausbildung entwickeln (Christensen et al., 2011).

2.3.1 Der Simulierte Klassenraum

Der Simulierte Klassenraum (SKR), welcher im Rahmen dieser Dissertation zum Simulierten Klassenraum Biologie (SKR^{Bio}) weiterentwickelt wurde, ist eine nicht-immersive Simulation, welche eine virtuelle Schulklasse mithilfe einer Kombination von Texten und Bildern darstellt. In dieser simulierten Klassenraumumgebung können die Studierenden in die Rolle der Lehrkraft schlüpfen und virtuelle Schülerinnen und Schüler zu einem bestimmten Thema unterrichten. Während einer virtuellen Unterrichtsstunde können Fragen aus einem Fragenmenü ausgewählt werden, worauf die virtuellen Schülerinnen und Schüler antworten. Hierbei können die virtuellen Schülerinnen und Schüler experimentell variiert werden, indem bei ihnen unterschiedliche Fähigkeitsprofile angelegt werden. Beispielsweise kann innerhalb der virtuellen Schülerschaft die Wahrscheinlichkeit richtiger Antworten oder die Beteiligung am Unterricht voreingestellt werden. Auf Basis gegebener Antworten der virtuellen Schülerinnen und Schüler oder der gezeigten Beteiligung am Unterricht können Einschätzungen durch die Lehramtsstudierenden vorgenommen werden. Die Übereinstimmung der Einschätzungen durch die Lehramtsstudierenden und der tatsächlich gezeigten Leistungen der virtuellen Schülerinnen und Schüler ergibt ein Maß für die Genauigkeit der Beurteilungen (Südkamp et al., 2008).

Anfänglich wurde der SKR als computergestütztes Messinstrument genutzt, um systematische Verzerrungen, beispielsweise der Tendenz zur Mitte oder der Einfluss von Stereotypen, in der Beurteilung virtueller Schülerleistungen durch die Studierenden zu untersuchen (Fiedler et al., 2002, 2007). Darauf aufbauend erweiterten Südkamp et al. (2008) den bestehenden SKR, um komplexere pädagogisch-psychologische Fragestellungen zu beantworten. Nun war es für die Forschenden möglich, eigene Fragestellungen und virtuelle Schülerantworten in das System zu integrieren, wodurch die diagnostischen Urteile von Lehramtsstudierenden differenzierter analysiert werden konnten. Außerdem erhöhte die Fokussierung auf unterschiedliche Unterrichtskontexte die ökologische Validität des SKR. In den vergangenen Jahren wurde diese simulierte Klassenraumumgebung eingesetzt, um spezifische Urteilsverzerrungen bei Lehramtsstudierenden experimentell nachzuweisen. So konnten innerhalb des SKR beispielsweise Stereotype gegenüber Minoritäten (Kaiser et al., 2016) oder der Big-Fish-Little-Pond-Effekt (Südkamp und Möller, 2009) nachgewiesen werden.

3 ZIELSETZUNG DER ARBEIT UND ZUSAMMENFASSUNG DER DURCHGEFÜHRTEN STUDIEN

Das Professionswissen angehender Biologielehrkräfte bildet die Grundlage für einen effektiven Unterricht und die zukünftigen Leistungen der Schülerinnen und Schüler. Empirische Studien lieferten Erkenntnisse, dass sich ein elaboriertes Professionswissen aus dem FW, FDW und PW zusammensetzt. Besonders im Zusammenhang mit dem Unterrichten der Evolutionstheorie im Biologieunterricht ist es essentiell, dass Lehrerinnen und Lehrer auf ein elaboriertes Professionswissen zurückgreifen können. Zahlreiche Studien belegen, dass vor allem beim Lehren und Lernen der Evolution zahlreiche Fehlvorstellungen existieren, welche den Aufbau des wissenschaftlich korrekten evolutionsbiologischen Konzepts behindern können. Dementsprechend müssen angehende Lehrerinnen und Lehrer der Biologie darauf vorbereitet werden, zukünftige Antworten der Schülerinnen und Schüler zur Evolution hinsichtlich der wissenschaftlichen Korrektheit beurteilen zu können, und gleichzeitig potenzielle spezifische Fehlvorstellungen zu identifizieren. Da in diesem Fall das professionelle Wissen in eine konkrete Handlung eingebettet wird, ist neben dem im akademischen Diskurs erlernten deklarativen Wissen auch prozedurales Wissen relevant. In der fachdidaktischen Forschung gibt es nur wenige Befunde darüber, inwiefern deklaratives und prozedurales Professionswissen miteinander interagieren und welche Instrumente sich anbieten, diese Wissenstypen adäquat zu erfassen. Hier bieten simulierte Klassenraumumgebungen eine Lösung an. Diese haben bereits in pädagogisch-psychologischen Studien gezeigt, dass durch die Interaktion der Lehramtsstudierenden innerhalb dieser Simulationen eine Handlungsnähe generiert werden kann, und somit das prozedurale Wissen adressiert wird.

Das Ziel der vorliegenden Dissertation ist es, bereits bestehende Arbeiten zum Professionswissen, zum Lehren und Lernen von Evolution sowie zu simulierten Klassenraumumgebungen aufzugreifen und weiterzuentwickeln. Dazu werden (1) evolutionsbiologische Inhalte, welche den Prozess der natürlichen Selektion betreffen, aufbereitet, um (2) diese in den Simulierten Klassenraum Biologie (SKR^{Bio}) zu integrieren, und somit für biologiedidaktische Fragestellungen anzuwenden, und (3) prozedurale biologiebezogene und generische Facetten des Professionswissens, sowie deren Zusammenhänge mit dem deklarativen Wissen, angehender Lehrerinnen und Lehrer zu untersuchen. In den folgenden Abschnitten wird eine kurze Übersicht über die drei durchgeführten Studien dieser Dissertation gegeben. Darüber hinaus gibt die Tabelle 3.1 einen Überblick über die wichtigsten Informationen jeder Studie.

3.1 Studie 1 (Kapitel 4): Measuring Biology Trainee Teachers' Professional Knowledge about Evolution – Introducing the Student Inventory

Mehrere Studien berichten über zahlreiche Fehlvorstellungen, welche bei der Erklärung des Prozesses der natürlichen Selektion geäußert werden. Bleiben diese Fehlvorstellungen unentdeckt, kann die Entwicklung wissenschaftlich korrekter Denkweisen behindert werden. Aktuell gibt es keine Befunde darüber, inwiefern Lehrerinnen und Lehrer im Vorbereitungsdienst (LiV) in der Lage sind, die Wissenschaftlichkeit sowie potenzielle spezifische Fehlvorstellungen zur natürlichen Selektion in Schülerleistungen zu diagnostizieren. Weiterhin ist aus der pädagogisch-psychologischen Forschung bekannt, dass bei der Leistungsbeurteilung neben leistungsrelevanten Informationen auch leistungsirrelevante Informationen berücksichtigt werden, welche keinen Einfluss auf die Leistungsbeurteilung haben dürften und Urteile systematisch verzerren. Es bleibt die Frage, ob diese in der pädagogisch-psychologischen Forschung beschriebenen Urteilsverzerrungen auch bei LiV der Biologie nachzuweisen sind. Die erste Studie adressiert diesen Forschungsbedarf, indem ein digitales Instrument – das Schülerinventar (SI) – eingesetzt wird, in welchem virtuelle Schülerklausuren zum Prozess der natürlichen Selektion integriert wurden. Das SI wurde von $N = 27$ LiV bearbeitet, wodurch anhand der Komposition der virtuellen Schülerklausuren innerhalb des SI Einblicke in verschiedene Facetten des prozeduralen evolutionsbezogenen FW und FDW sowie des PW generiert werden konnten. Außerdem wurden mithilfe eines Fragebogens zur Evolution, welcher deklaratives Wissen der LiV erfasste, erste Hinweise auf den Zusammenhang zwischen deklarativen und prozeduralen Facetten des Professionwissens von LiV der Biologie untersucht.

3.2 Studie 2 (Kapitel 5): Simulated Classroom Biology (SCR^{Bio}) - a Simulated Classroom Environment for Assessing the Action-Oriented Professional Knowledge of Pre-Service Teachers about Evolution

Die Befunde der ersten Studie deuten darauf hin, dass die entwickelten virtuellen Schülerklausuren zur natürlichen Selektion den konzeptionellen Rahmen bieten, um biologiebezogene Facetten des Professionwissens zu erfassen. Dementsprechend dienen die Ergebnisse dieser Studie als Grundlage der weiteren Entwicklung des evolutionsbiologischen Inhalts für den SKR^{Bio}. Aus der Forschung ist bereits bekannt, dass simulierte Klassenraumumgebungen die Möglichkeit bieten, prozedurales Wissen vor allem im Bereich des PW zu messen. Es fehlen validierte simulierte

Klassenraumumgebungen, welche darüber hinaus in der Lage sind, verlässliche und objektive Ergebnisse zum prozeduralen Professionswissen zu liefern. In der zweiten Studie wurde daher der SKR^{Bio} mit $N = 76$ Biologie-Lehramtsstudierenden anhand von verschiedenen Evidenzkriterien für die valide Testwertinterpretation überprüft, um Evidenz zu generieren, ob prozedurales fachdidaktisches Wissen valide erfasst werden kann.

3.3 Studie 3 (Kapitel 6): Der Simulierte Klassenraum Biologie: Erfassung deklarativen und prozeduralen Wissens bei Lehramtsstudierenden der Biologie

Während in der Lehrerprofessionalisierungsforschung zahlreiche Studien durchgeführt wurden, welche intensiv die verschiedenen Wissensbereiche (FW, FDW und PW) des Professionswissen angehender Lehrkräfte untersuchten, wurde eine explizite Berücksichtigung des deklarativen und prozeduralen Wissens lange vernachlässigt. Dabei scheint es in der Forschung Hinweise zu geben, dass deklaratives Wissen der Ausgangspunkt für die Entwicklung prozeduralen Wissens darstellt, welches vor allem das Handeln einer Lehrperson im Unterricht determiniert. Besonders die Frage nach dem Zusammenhang von deklarativem und prozeduralem biologiebezogenen Professionswissen konnte durch bisherige Forschung nicht hinreichend beantwortet werden. Vor diesem Hintergrund und nach erfolgreicher Validierung des SKR^{Bio} in der zweiten Studie wird in der dritten Studie eine Kombination zweier Instrumente verwendet. Diese sind ein Fragebogen zur Evolution, welcher deklaratives Wissen misst und der SKR^{Bio}, der prozedurales Wissen erfasst, um mit $N = 51$ Lehramtsstudierenden der Biologie zu untersuchen, inwiefern deklaratives und prozedurales biologiebezogenes Professionswissen zur Evolution zusammenhängen.

ZIELSETZUNG UND ÜBERSICHT DER STUDIEN

Tabelle 3.1 Übersicht über die durchgeführten Studien

Studie 1 (Kapitel 4)	
Veröffentlichung / Manuskript	Fischer, J., Jansen, T., Möller, J., & Harms, U. (2021). Measuring Biology Trainee Teachers' Professional Knowledge about Evolution – Introducing the Student Inventory. <i>Evolution: Education and Outreach</i> , 14(1), 1–16.
Biologischer und biologiedidaktischer Inhaltsbereich	<ul style="list-style-type: none"> – Prozess der natürlichen Selektion des Birkenspanners – Wissen über Schülerfehlvorstellungen (Fehlvorstellungskategorie: anthropomorph und teleologisch)
Zielsetzung	<ul style="list-style-type: none"> – Untersuchung unterschiedlicher Facetten des biologiebezogenen und generischen Professionswissens sowie erste Hinweise über den Zusammenhang zwischen deklarativem und prozeduralem Wissen.
Forschungsfragen	<ol style="list-style-type: none"> 1. Inwieweit sind LiV in der Lage, wissenschaftlich korrekte von wissenschaftlich inkorrekten schriftlichen Schülerantworten zum evolutionären Prozess der natürlichen Selektion zu unterscheiden? 2. Inwieweit sind LiV in der Lage, die Fehlvorstellungskategorie (anthropomorph oder teleologisch) in den schriftlichen Antworten der virtuellen Schüler über den evolutionären Prozess der natürlichen Selektion zu diagnostizieren? 3. Inwieweit beeinflussen leistungsirrelevante Informationen (frühere Leistungen) die Diagnose späterer Leistungen (leistungsrelevante Informationen) und führen zu Beurteilungsfehlern bei LiV? 4. Welche ersten Hinweise auf den Zusammenhang zwischen dem deklarativen Wissen und dem prozeduralen Wissen der LiV können beobachtet werden?
Design und Stichprobe	<ul style="list-style-type: none"> – Deskriptive Querschnittsstudie mit einem 2x3 Versuchsplan – $N = 27$ Lehrerinnen und Lehrer im Vorbereitungsdienst der Biologie
Instrumente	Das Schülerinventar

Tabelle 3.1 Fortsetzung

Studie 2 (Kapitel 5)	
Veröffentlichung / Manuskript	Fischer, J., Machts, N., Bruckermann, T., Möller, J., & Harms, U. (eingereicht 2021). Simulated Classroom Biology (SCR ^{Bio}) - a Simulated Classroom Environment for Assessing the Action-Oriented Professional Knowledge of Pre-Service Teachers about Evolution. <i>Journal of Science Teacher Education</i> .
Biologischer und biologiedidaktischer Inhaltsbereich	<ul style="list-style-type: none"> – Prozess der natürlichen Selektion unterschiedlicher Organismengruppen (Bakterien, Pflanzen, Tiere, Mensch) – Wissen über Schülerfehlvorstellungen (Fehlvorstellungskategorie: anthropomorph, lamarckistisch und teleologisch)
Zielsetzung	<ul style="list-style-type: none"> – Erprobung und Validierung des Simulierten Klassenraums Biologie zur validen Testwertinterpretation prozeduralen fachdidaktischen Wissens
Forschungsfragen	<ol style="list-style-type: none"> 1. Inwieweit stützen die Expertenurteile die Evidenz für die Validität der entwickelten virtuellen Schülerantworten in Bezug auf die spezifischen Fehlvorstellungskategorien? (Evidenz basierend auf dem Testinhalt) 2. Inwieweit ist der SKR^{Bio} sensibel für gegebene biologiedidaktische Informationen in Bezug auf gezeigtes biologiedidaktisches sowie pädagogisches Wissen der Lehramtsstudierenden? (Evidenz basierend auf Gruppenvergleichen)
Design und Stichprobe	<ul style="list-style-type: none"> – Experimentelle Validierungsstudie; untersuchte Evidenzkriterien für die valide Testwertinterpretation: (1) Evidenz basierend auf dem Testinhalt, (2) Evidenz basierend auf Gruppenvergleichen (Querschnittsstudie) – $N = 76$ Lehramtsstudierende der Biologie (Bachelor und Master Studium)
Instrumente	Der Simulierte Klassenraum Biologie

ZIELSETZUNG UND ÜBERSICHT DER STUDIEN

Tabelle 3.1 Fortsetzung

Studie 3 (Kapitel 6)	
Veröffentlichung / Manuskript	Fischer, J., Machts, N., Möller, J., & Harms, U. (major revisions 2021). Der Simulierte Klassenraum Biologie: Erfassung deklarativen und prozeduralen Wissens bei Lehramtsstudierenden der Biologie. <i>Zeitschrift für die Didaktik der Naturwissenschaften</i> .
Biologischer und biologiedidaktischer Inhaltsbereich	<ul style="list-style-type: none"> – Fragebogen: Makroevolution, Mikroevolution, Erklärungsmodelle von Vielfalt – Wissen über Schülerfehlvorstellungen und Wissen über Instruktionsstrategien – SKR^{Bio}: Prozess der natürlichen Selektion unterschiedlicher Organismengruppen (Bakterien, Pflanzen, Tiere, Mensch) – Wissen über Schülerfehlvorstellungen (Fehlvorstellungskategorie: anthropomorph, lamarckistisch und teleologisch)
Zielsetzung	– Zusammenhang zwischen deklarativen und prozeduralen Facetten des Professionswissens zur Evolution
Forschungsfragen	Inwieweit hängen das deklarative und prozedurale Wissen zur Evolution der Lehramtsstudierenden zusammen?
Design und Stichprobe	<ul style="list-style-type: none"> – Explanative Querschnittsstudie – $N = 51$ Lehramtsstudierende der Biologie (Master-Studium)
Instrumente	<p>ProwiE-Fragebogen zum fachlichen und fachdidaktischen Wissen zur Evolution von Großschedl et al. (2015b)</p> <p>Der Simulierte Klassenraum Biologie</p>

4 STUDIE 1: MEASURING BIOLOGY TRAINEE TEACHERS' PROFESSIONAL KNOWLEDGE ABOUT EVOLUTION – INTRODUCING THE STUDENT INVENTORY³

Abstract:

Background

To teach evolution efficiently teachers must be able to diagnose their students' ideas and understanding of the phylogeny of organisms. This encompasses different facets of content-specific professional knowledge, that is, knowledge about core ideas and theories, as well as knowledge about respective misconceptions. However, as findings from the field of psychology have shown, diagnostic activities comprise a further facet, namely, teachers' judgment accuracy. This refers to the question of whether achievement-irrelevant information about the student influences teachers' diagnoses. Against this background we conducted a study (1) to assess trainee teachers' abilities to diagnose (a) the scientific correctness of students' written answers, (b) students' misconceptions about evolution, and (2) to investigate the interplay of evolution-specific and generic facets of professional knowledge during the diagnosis. For this purpose, we applied a digital instrument, the Student Inventory (SI). Using this instrument, the trainee teachers ($N = 27$) first diagnosed written answers ($N = 6$) from virtual students regarding their scientific correctness and regarding students' misconceptions about the natural selection of the peppered moth. Second, to test for judgment accuracy, the trainee teachers received – via the SI – achievement-irrelevant information about each virtual student, that is, the previous result of a multiple-choice questionnaire about evolution, before diagnosing the written answers.

Results

The trainee teachers were able to distinguish between scientifically correct (90.8%) and scientifically incorrect (91.7%) written answers. Trainee teachers faced problems when diagnosing specific misconceptions categories. Anthropomorphic misconceptions were diagnosed significantly more often (61.1%) than teleological misconceptions (27.8%). The achievement-irrelevant information influenced the trainee teachers' assessment of written answers ($F[1,26] = 5.94, p < .022, \eta^2 = .186$) as they scored the written answers higher if the performance in the questionnaire was good and vice versa.

³ Fischer, J., Jansen, T., Möller, J., & Harms, U. (2021). Measuring Biology Trainee Teachers' Professional Knowledge about Evolution – Introducing the Student Inventory. *Evolution: Education and Outreach*, 14(1), 1–16.

Conclusion

The findings indicate that the diagnosis is easier or more difficult depending on the particular misconception category. However, the findings also reveal that, besides the evolution-specific facets of professional knowledge, generic facets interrelate with the quality of the diagnosis result. We conclude from these findings that an integration of evolution-specific and generic knowledge into the education of biology teachers is critical.

4.1 Introduction

The mission of evolution education is to foster accurate mental models of the mechanisms of evolutionary theory – the overarching framework of the life sciences – and to introduce an appreciation of the centrality of this framework for a scientific understanding of the living world. However, students throughout the educational hierarchy, the public, and even science teachers lack an understanding of the relevant principles and concepts of evolutionary change (e.g., Nadelson and Sinatra 2009; Nehm and Reilly 2007). Many of them also resist accepting the theory of evolution as the best scientific explanation for the similarities among organisms, for biological diversity, and for various features and processes in the living world (Berkman and Plutzer 2011). The factors influencing understanding are diverse and include influences of alternative conceptions (e.g., Kampourakis and Nehm 2014), mismatches of everyday language and scientific terminology (Rector et al. 2013), selection of model organisms and traits (i.e., animal vs. plant, trait gain vs. trait loss; Großschedl et al. 2018), students' thinking dispositions (Athanasiou and Papadopoulou 2012), feeling of certainty (Ha et al. 2012), and students' religious views (e.g., Allmon 2011; Athanasiou and Papadopoulou 2012). Thus, the correct teaching of the evolutionary theory by biology teachers is highly important for students, as it acts as a central link between different concepts and highlights the similarities in the complexity of biological concepts (Tibell and Harms 2017).

The teacher is one of the most important determinants of students' performance (e.g., Mahler et al. 2017). Teachers make a difference in the achievement of their students in science classrooms. Many studies have shown that teachers' professional knowledge is the key factor for teaching (e.g., Abell 2007; Kunter et al. 2013). Based on Shulman's (1987) taxonomy, professional knowledge in educational sciences can be differentiated into three main facets: content knowledge (CK), pedagogical content knowledge (PCK), and pedagogical knowledge (PK). In the context of evolution education, for example, CK is necessary to identify key ideas and principles about evolutionary mechanisms such as natural selection (e.g., Großschedl et al. 2015), PCK to diagnose prominent student misconceptions (e.g., Ziadie and Andrews 2018), and PK to make accurate judgments about the performance shown (e.g., Schrader 2006). Thus, these domains of professional knowledge and the corresponding facets play an important role in the diagnostic activities of teachers, for example, in the assessment of students' statements or written performance (e.g., Helmke et al. 2004). Without a comprehensive diagnosis of existing misconceptions, a subsequent individualized support of the students to enable them to succeed in achieving an elaborated understanding of evolution is not possible.

The aim of the present study was to investigate trainee teachers' biology-specific and generic facets of professional knowledge. The focus was on knowledge regarding

evolutionary theory, specifically the process of natural selection. In an effort to gain insights into biology teachers' diagnostic activities based on their professional knowledge, we further developed the digital survey instrument – the Student Inventory (SI) - of Kaiser et al. (2015). Furthermore, we gained first indications about the relationship between the declarative and procedural knowledge of the biology trainee teachers. In the following sections, we strongly focus on the teaching and learning of evolution, focusing on prominent student misconceptions related to natural selection as well as on the biology trainee teachers' professional knowledge.

4.2 Theoretical Background

4.2.1 Teaching and Learning Evolution

Evolution is the central comprehensive explanatory framework not only in biology but also in all of the life sciences (e.g., German National Academy of Sciences Leopoldina 2017; National Research Council 2012). It has the power to explain the diversity of life and to foster the understanding of how and why populations change over time (e.g., Sickel and Friedrichsen 2013). Therefore, the theory of evolution is part of numerous science education curricula and standards in many countries (e.g., Germany: Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States in the Federal Republic of Germany [KMK] 2005; USA: Next Generation Science Standards [NGSS] 2013). Evolutionary processes are the basis for all topics in the field of biology and form the foundation for a conceptual understanding of the life sciences (e.g., Anderson et al. 2002; Basel et al. 2014; Bishop and Anderson 1990; Furtak 2012; Opfer et al. 2012; Zabel and Gropengießer 2011). Despite the complexity of the theory of evolution, the core ideas of evolution can be summarized in a few sentences. Biological evolution can be defined at the lowest level (microevolution) as changes in the allele frequency of a population. These changes and the resulting genetic variation are caused by recombination and mutation. Variability is the basis of the core mechanism of evolution – natural selection, which describes the adaptation within species. Those individuals who are genetically better adapted to environmental conditions are more likely to survive as well as to have a higher reproduction rate. In the next generation, these beneficial traits occur more frequently and the population adapts to the environmental conditions (e.g., Andersson and Wallin 2006). Although there are differences in the number of relevant key concepts (e.g., Anderson et al. 2002; Moharrerri et al. 2014; Nehm and Schonfeld 2008), the concepts of variation, inheritance, and selection provide a sufficient explanation of evolutionary change through the process of natural selection (Endler 1986; Mayr 1982).

However, the complexity of the evolutionary theory leads to great problems in teaching and learning. Research has identified evolution as a very challenging component of the science curriculum and indicates that traditional teaching approaches are ineffective in transforming students' misconceptions into scientifically adequate ways of thinking (e.g., Basel et al. 2013; Bishop and Anderson 1990; Gregory 2009; Kampourakis and Zogza 2008; McVaugh et al. 2011; Opfer et al. 2012; Pazza et al. 2010). A major impediment to learning is the numerous misconceptions that exist with regard to the theory of evolution. Misconceptions are understood to be students' ideas and thoughts that are incompatible with scientific knowledge (Yip 1998). Students' misconceptions are dependent on their experiences, the language used in daily communication, the lack of CK of their teachers, and the textbooks they use (e.g., King 2010; Nehm et al. 2009). Misconceptions about natural phenomena arise at an early age and even before school education, when children explore their physical and social environments (e.g., Beardsley 2004; Bruckermann et al. 2020; Driver 1988; Evans 2000). These misconceptions are used to explain evolutionary mechanisms. Consequently, misconceptions not only are found at the level of young children but are also anchored in the minds of high school students (Demastes et al. 1995), biology majors (e.g., Dagher and BouJaoude 1997; Nehm and Reilly 2007; Nehm and Schonfeld 2008), medicine students (Brumby 1984), and science teachers (Nehm and Schonfeld 2007). The fact that the principles of evolutionary biology are widely misunderstood by students as well as by large parts of the public has motivated educators and researchers to focus on identifying evolutionary misconceptions and finding instructional strategies to overcome these (for an overview: on misconceptions, see Gregory 2009; on teaching strategies, see Ziadie and Andrews 2018; Harms and Reiss 2019).

Evolutionary misconceptions are manifold and refer to different evolutionary mechanisms. Evans (2000) observed that the process of the origin of species is seen as a spontaneous event regardless of the evolutionary processes involved. This is expressed in numerous creationist or religious ideas that explain that God or another creator is responsible for the origin of species (e.g., Basel et al. 2014; Berti et al., 2010; Billingsley et al. 2016; Großschedl et al. 2014; Rissler et al. 2014; Yasri and Mancy 2014). Here, evolutionary biological knowledge is particularly important, as international studies have shown that sophisticated knowledge about evolution can positively influence the acceptance of the theory of evolution (e.g., Barnes et al. 2017; Deniz et al. 2008; Fiedler et al. 2019; Ha et al. 2015). Another set of problems results from the fact that many evolutionary concepts appear to be counterintuitive to students (e.g., Tibell and Harms 2017). In this context, misconceptions can arise in the field of phylogeny, where the depth time of evolutionary processes is not understood correctly (e.g., van Dijk and Kattmann 2010), or the origin of related species leads to problems for students when the concept of

STUDIE 1

the last common ancestor cannot be discerned (e.g., Baum et al. 2005; Catley et al. 2013; Gregory 2008; Phillips et al. 2012).

Natural selection is the key mechanism of evolutionary change that leads to features adapting to new environmental conditions. The concept of natural selection is widely accepted by biologists today and can be briefly summarized: Species are adapted to their environment because individuals with the most suitable traits for that environment have a higher probability of survival and pass these traits on to their offspring. Over time, this leads to changes in the frequency of the hereditary traits of populations (Mayr 1982). This simple explanation of natural selection suggests that it would be easy to communicate, but years of research have shown that it is one of the most difficult topics to teach in biology (e.g., Bishop and Anderson 1990; Nehm and Reilly 2007). In connection with the process of natural selection, numerous misconceptions that inhibit students' understanding are described in the literature (e.g., Bishop and Anderson 1990; Ferrari and Chi 1998; Gregory 2009; Nehm et al. 2009; Shtulman 2006). These include, in particular, anthropomorphic and teleological misconceptions. Problems arise when students transfer human thinking, including emotion, motivation, and reasoning, to non-human organisms such as animals or plants. These anthropomorphic beliefs are based on the conception that the change of a trait is the result of an intentional and purposeful action performed by the individual to cope with new environmental conditions (e.g., Byrne et al. 2009; Demastes et al. 1995; Gregory 2009; Kallery and Psillos 2004; Sinatra et al. 2008; Tamir and Zohar 1991). An example of an anthropomorphic explanation is that the eagle's good eyes have developed because the eagles thought that good eyes would help them to spot the mouse from a far distance (Neubrand 2017). Here, the development of the eyes is directed solely by the individual, who judges the characteristic to be beneficial. This trait development is a singular event and does not refer to any evolutionary mechanisms. A further conceptual bias related to anthropomorphism is teleology, in which the environment itself causes traits to change over time. Teleological misconceptions always follow a "start-finish scheme" with an unchangeable final result (Stover and Mabry 2007). Here, the development of a trait is target-oriented and purposeful (e.g., Alters and Nelson 2002; Andrews et al. 2011; Beardsley 2004; Bishop and Anderson 1990; Kampourakis and Zogza 2008; Nehm et al. 2009; Nehm and Reilly 2007; Nehm and Schonfeld 2008; Settlage 1994; Sinatra et al. 2008). To explain the good eyes of the eagle, students could say that they have evolved in order to give the eagle an advantage when hunting (Neubrand 2017). However, the dynamics of the adaptation process of living organisms are far more complex, and underlying concepts of evolution such as the influence of randomness and probability are completely ignored in these misconceptions (e.g., Fiedler et al. 2017, 2018; Garvin-Doxas and Klymkowsky 2008).

The use of the goal- and purpose-oriented explanation of evolution is a natural human tendency and is intuitively based on humans' own personal experiences in goal- and problem-oriented thinking (Gregory 2009). This process is reinforced by the fact that pupils are often asked in science classes to explain natural phenomena causally (Olander 2012). Additionally, this tendency is supported by everyday language but also by scientific language (e.g., Alters and Nelson 2002; Nehm et al. 2010). Terms such as selection and adaptation suggest that these are directed processes that can in fact be viewed as beneficial under the current environment (e.g., Baalman et al. 2004; Gregory 2009). However, the goal or purpose is not the determining factor for the development of a trait instead, evolutionary biological mechanisms such as variability, selection, and inheritance are the determining factors (e.g., Godfrey-Smith 2007; Tibell and Harms 2017).

Frequently, there are problems in distinguishing between the individual and population level in evolutionary processes. The process of natural selection is based on individual traits and their interrelation with the environment. Finally, it is genetic variability that causes the differences in the phenotype. Individuals in a population therefore exhibit morphological, physiological, and behavioral differences, which can manifest themselves through generations in a population (Andersson and Wallin 2006). If the precondition that adaptation takes place on an individual level is ignored, an essentialist view can result. The essentialist misconception is characterized by the assumption that members can be assigned to a category that has an underlying "true nature" that is permanent and heritable. This true nature gives these members their basic identity (Evans 2000; Shtulman 2006). Here, differences between the evolutionary processes of populations (i.e., between-category differences) are overestimated, whereas variability at the individual level (within-category differences) is underestimated, which poses a threat to the understanding of the evolutionary theory (Opfer et al. 2012).

Learning difficulties regarding the topic of evolution have been shown in many studies (e.g., Nehm and Reilly 2007; Nehm and Schonfeld 2007; Wandersee et al. 1995) and this finding has spurred researchers to focus on identifying and addressing common misconceptions (Anderson et al. 2002). Thus, the overarching goal of biology teaching is to support students to acquire conceptually and biologically correct knowledge about evolution and to prevent misconceptions (Gregory 2009). Here, several facets of the trainee teachers' professional knowledge, which forms the basis for diagnostic activities, are necessary to evaluate scientific correctness and are essential to identify students' difficulties in understanding evolution (i.e., misconceptions) and to make adequate interventions.

4.2.2 Biology Teachers' Professional Knowledge

Every day, biology teachers are confronted with diagnostic activities in the classroom. These include, for example, assessing the correctness of student answers during lessons or evaluating written performance, as shown in exams (Förtsch et al. 2018). Within the domain of professional knowledge, CK, PCK, and PK are relevant in the diagnosis of student performance (e.g., Brunner et al. 2011; Helmke et al. 2004; Kunter et al. 2013). Both CK and PCK are primarily described as content-specific facets, while PK can be considered as content-independent (e.g., Förtsch et al. 2018).

CK in general addresses the knowledge about facts and terms as well as conceptual understanding (Shulman 1986). In the context of evolution, CK primarily comprises the knowledge about the key ideas and principles of evolution. In addition, biology-specific CK includes the knowledge to determine validity within the domain (i.e., knowledge of research methods) and the knowledge about the nature of science (e.g., Großschedl et al. 2015). Several studies have shown that elaborated CK is essential for effective teaching (e.g., Baumert et al. 2010; Friedrichsen et al. 2009), but CK alone is not sufficient to enable teachers to perform diagnostic activities that lead to adaptive teaching and interventions in learning (e.g., Abell 2007; Baumert et al. 2010; Förtsch et al. 2018). CK is an important prerequisite for the development of PCK, which was defined by Shulman (1987) as a synthesis of content and pedagogy, and goes beyond subject matter knowledge. This knowledge domain is required to make the subject matter understandable. Shulman's model (1987) describes at least two facets of PCK, the knowledge about students' conceptions and preconceptions, and the knowledge about strategies to overcome them. Numerous research groups agreed with this initial conceptualization and defined the knowledge about students' understanding and the knowledge about instructional strategies for teaching as the most important facets of PCK (e.g., Förtsch et al. 2018; Grossman 1990; Hill et al. 2008; Lee and Luft 2008; Mahler et al. 2017; Park and Oliver 2008). Knowledge about student misconceptions includes knowledge about the context in which student misconceptions occur, the context-specific categories of misconceptions, and the extent to which these misconceptions can impede the learning of scientific concepts. By anticipating these misconceptions, a teacher can plan questions to reveal this thinking and to teach in such a way that will help students to develop scientifically adequate ideas about natural selection (Ziadie and Andrews 2018). The knowledge about instructional strategies comprises knowledge on how to integrate the representation of subject matter and how to address specific learning difficulties (Großschedl et al. 2015; Hill et al. 2008; Lee and Luft 2008). Additionally, other facets of PCK have been introduced in the past, such as knowledge of the curriculum (e.g., Tamir 1988; Ziadie and Andrews 2018), knowledge of assessment methods (e.g., Hashweh 2005; Magnusson et al. 1999; Ziadie and Andrews 2018), knowledge about models (Tepner et al. 2012), or knowledge of teaching resources

(Lee and Luft 2008). To diagnose whether a student has already developed a scientific concept in evolution, the teacher needs knowledge about the key ideas and principles of evolution as a facet of CK. If the student holds a misconception, the first facet of PCK, the knowledge about student understanding, is relevant. The teacher must assess the quality of the student's understanding, that is, which type of misconception is present (Förtsch et al. 2018).

In comparison to CK and PCK, facets of PK transcend the content-related areas and focus on knowledge about learning strategies, knowledge about effective classroom management, and knowledge about judgment accuracy (e.g., Brunner et al. 2011; Kunter et al. 2013). The latter refers to the ability to assess individuals appropriately (Schrader 2006). Previous research on the diagnostic competence (here: knowledge about judgment accuracy; facet of PK) of teachers has shown that teachers' judgments are influenced by judgment errors, such as the halo effect, and affect the judgment accuracy (Vögelin et al. 2019). A halo effect occurs when one feature affects the judgment of another independent feature. The halo effect refers to the tendency to form an overall impression based on a prominent, dominant feature, which prevents the teacher from distinguishing between different features of performance assessment (e.g., Borman 1975; Murphy and Reynolds 1988).

Regardless of this differentiation between CK, PCK, and PK, the dichotomous classification of teacher professional knowledge into declarative ("knowledge that") and procedural ("knowledge how") knowledge has been established based on psychological approaches (e.g., Fenstermacher 1994; König et al. 2014). Declarative professional knowledge comprises factual knowledge that is accessible (explicit) to consciousness and is acquired primarily in academic discourse (Baumert and Kunter 2013). Procedural professional knowledge comprises action-oriented knowledge, which is often implicit and therefore difficult to verbalize. Through systematic practice and contextualization, procedural knowledge may develop from declarative knowledge (e.g., Schneider and Stern 2010). However, the two types of knowledge can also be unconnected or even contradictory (Shulman 1986). This reveals a problem in teacher education, namely, that the declarative knowledge acquired in teacher education often remains tacit and is not transformed into procedural knowledge (e.g., Blömeke et al. 2008; Renkl 1996). Accordingly, declarative knowledge cannot be retrieved and applied in real classroom situations. Instead, unreflected beliefs based on personal experience as a student, trainee, or teacher often determine practical action.

4.3 Research Questions

The overall goal of this study, in addition to investigating the content-related facets of professional knowledge on evolution (i.e., CK and PCK), was to capture generic knowledge (i.e., PK) in order to gain deeper insights into the complex diagnostic activities of biology trainee teachers. Therefore, the selected facets of professional knowledge were operationalized with virtual student exams and transferred into a digital instrument - the Student Inventory (SI). The SI allowed us to experimentally vary different information within a virtual student exam and, based on this, to analyze trainee teachers' diagnostic activities in different facets of the trainee teachers' generic as well as biology-specific professional knowledge. Another strength of the SI is that it ensured that each trainee teacher received the accurate variation in a standardized way, resulting in high implementation fidelity. We integrated virtual student exams on evolution into the SI, which each presented a multiple-choice performance and a written answer. In order to assess the students' written answers, biology trainee teachers had to apply their knowledge about the core ideas and principles of evolution (facet of CK) and their knowledge about student understanding (facet of PCK) to assess scientific quality as well as potentially existing misconceptions. To arrive at adequate diagnoses within the virtual student exams, that is, to include only relevant information in the assessment, trainee teachers needed their knowledge about judgment accuracy (facet of PK). The facets of professional knowledge needed to assess the virtual student exams were conceptualized as procedural knowledge, because the trainee teachers had to apply their knowledge in a specific action-related situation, that is, during their assessment of the virtual student exams (Förtsch et al. 2018; Kaiser et al. 2015). Additionally, we examined whether declarative knowledge, assessed in a short questionnaire about knowledge of evolution, influenced the procedural knowledge surveyed in the SI (see Figure 4.1).

One facet of the CK knowledge domain is the knowledge about the core ideas and principles of evolution (Großschedl et al. 2015). This facet of CK had to be applied by the trainee teachers in order to assess the scientific quality of the students' written answers, which meant differentiating between scientifically correct and scientifically incorrect explanations.

RQ 1: To what extent are trainee teachers able to distinguish scientifically correct from scientifically incorrect students' written answers on the evolutionary process of natural selection in the SI? (i.e., knowledge about the core ideas and principles of evolution – facet of CK; procedural knowledge)

Knowledge about student understanding is, according to many studies in science education, a central facet of PCK (e.g., Förtsch et al. 2018; Hill et al. 2008; Lee and Luft 2008; Mahler et al. 2017; Park and Oliver 2008). This facet of PCK enables the trainee

teachers to identify specific misconceptions about evolution in the virtual students' written answers.

RQ 2: To what extent are trainee teachers able to diagnose the misconception category (i.e., anthropomorphic or teleological) in the students' written answers on the evolutionary process of natural selection in the SI? (i.e., knowledge about student understanding – facet of PCK; procedural knowledge)

Educational psychological research has described numerous judgment errors made by teachers that contribute significantly to the distortion of judgment accuracy. The knowledge about judgment accuracy (facet of PK) is therefore necessary in order for teachers to make accurate judgments (e.g., Jansen et al. 2019, 2021; Kaiser et al. 2015; Schrader 2006; Vögelin et al. 2019). Studies have already shown that teachers have problems in assessing relevant performance without including previously shown performance, which should actually be assessed independently (Malouff and Thorsteinsson 2016; Oudman et al. 2018). Thus, the quality of an answer in a previous task within an exam can have an impact on the assessment of the quality of a subsequent answer. However, only achievement-relevant information should be considered, that is, in the present study, performance in a student's written answer on the natural selection of the peppered moth. Any influence of a previous performance in a multiple-choice test on evolution can be seen as causing bias as the previous performance is achievement-irrelevant information.

RQ 3: To what extent does achievement-irrelevant information (previous performance) influence the diagnosis of subsequent performance (achievement-relevant information) and lead to judgment errors among trainee teachers? (i.e., knowledge about judgment accuracy – facet of PK; procedural knowledge)

In psychological approaches, two types of knowledge have been identified, which differ in their applicability to different situations (e.g., Fenstermacher 1994; König et al. 2014). Knowledge that is necessary for answering questions in a questionnaire is operationalized as factual knowledge and is thus assigned to declarative knowledge (Baumert and Kunter 2013). However, if knowledge is involved in a specific context of action, which is the case when assessing students' written answers in the SI, it is operationalized as procedural knowledge (Schneider and Stern 2010).

RQ 4: Which first indications of the interrelationship can be observed between the trainee teachers' declarative knowledge (i.e., facets of professional knowledge, which are surveyed in a questionnaire on evolution) and their procedural knowledge (i.e., facets of professional knowledge surveyed in the SI)?

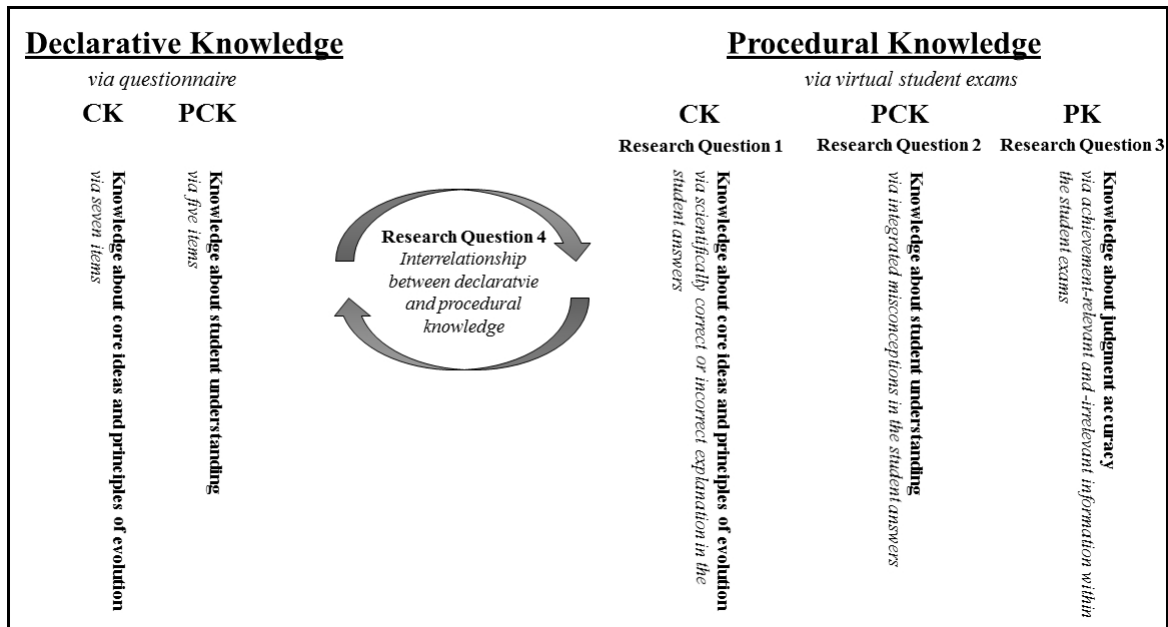


Abbildung 4.1 An overview of the investigated knowledge types, knowledge domains, and knowledge facets, as well as the operationalization within the SI and related research questions.

4.4 Method

4.4.1 Sample

The SI was completed by 27 in-service trainee teachers ($N = 27$; 22% male). In Germany, the teacher education program is divided into university education (i.e., bachelor's and master's degree; first state exam) and in-service training (second state exam). The university teacher education encompasses three and a half to five years and focuses on the development of CK, PCK, and PK. Within this time, there are short practical phases in schools, which last between two and five months (KMK, 2014). In our study, the trainee teachers had already attended lectures and courses that explicitly teach CK, that is, knowledge about the core ideas and principles of evolution, PCK, that is, knowledge about student understanding, and PK, that is, knowledge about judgment accuracy. The in-service training takes 18 months to two years and encompasses the teaching of a regular school class, which is guided by mentors (Neumann et al. 2017). All of the in-service trainee teachers in our sample aspired to a teaching qualification for academic track secondary schools (i.e., Gymnasium) in the federal state of Schleswig-Holstein (Germany) and studied

biology as one of their teaching subjects. On average, the trainee teachers were 29.3 ($SD = 4.8$) years old.

4.4.2 The Student Inventory (SI)

The SI is a digital instrument that can be used with a web browser and allows a split-screen on the PC monitor. This provides a multitasking function, where, for example, virtual student exams can be read on one side and an evaluation of the student exams can simultaneously be made on the other side. Hereby, the SI differs substantially from a regular questionnaire. The SI was developed by Kaiser et al. (2015) and was initially used exclusively for research in the field of educational psychology to measure the judgment accuracy of pre-service, trainee, or in-service teachers with regard to the assessment of student exams or separate tasks in mathematics. Therefore, the SI systematically combines achievement-relevant information with achievement-irrelevant information within student exams and investigates how these types of information influence the judgment accuracy. For example, if student performance is to be accurately assessed, oral and written performances shown in the lesson may be considered relevant to achievement. Other information should not be included in an accurate performance judgment and is therefore described as being irrelevant to achievement. Many studies have already examined the influence of student characteristics that are not relevant to judgment on teachers' judgments of student performance (e.g., Schrader and Helmke 1990; Ritts et al. 1992; Ready and Wright 2011). Kaiser et al. (2015) used the SI to investigate the judgment accuracy of trainee teachers; the teachers gave the virtual students simple mathematics exercises (addition, subtraction, multiplication, and division) and received either correct or incorrect answers (information relevant to achievement). Additionally, the trainee teachers received achievement-irrelevant information about each virtual student (grade in a German test, intelligence, self-concept, family background, and gender). Kaiser et al. (2015) were able to show that the achievement-irrelevant information German grade, intelligence, and gender (female) had a positive effect on the assessment of mathematics performance and thus biased the accuracy of judgments. Based on the same theoretical framework, a study by Jansen et al. (2019, 2021) that used the SI with pre-service teachers investigated, among other things, the extent to which students' gender or an immigrant background (achievement-irrelevant information) influenced the assessment of students' English essays. Experimental variation in achievement-relevant (essay quality) and achievement-irrelevant (gender, immigrant background) information showed no effect on the assessment accuracy of the English essays. In both studies, it was the teachers' task to diagnose and evaluate the achievement-relevant information. The comparison of the real performance of the students with the performance diagnosed by the teachers provided information on whether an accurate judgment was made or a bias in the judgment occurred due to the achievement-irrelevant information (Kaiser et al. 2017). In the above-mentioned studies,

STUDIE 1

the construct to be investigated was the knowledge about judgment accuracy, whereby initial aspects of content (i.e., professional judgment of English essays) were also considered (Vögelin et al. 2018).

We used the SI on the diagnostic competence of pre-service, trainee, and in-service teachers, which was developed by Kaiser et al. (2015) and used to investigate educational psychological research questions (Jansen et al. 2019, 2021). We adapted it for our purposes to measure the generic knowledge about judgment accuracy (i.e., diagnosis of achievement-relevant information; facet of PK) and, simultaneously, biology-specific facets of professional knowledge such as the knowledge about the core ideas and principles of evolution (i.e., diagnosis of scientific correctness; facet of CK) and the knowledge about student understanding (i.e., diagnosis of specific misconception categories; facet of PCK). Similar to the psychological studies, we also experimentally varied achievement-relevant with achievement-irrelevant information within virtual student exams in the SI. As achievement-relevant information, students' written answers on the natural selection of the peppered moth were integrated. The previous performance in a multiple-choice test on evolution was presented as achievement-irrelevant information. The trainee teachers' task was to assess the written answers of the virtual students without being influenced by the multiple-choice test performance previously achieved by the students.

As achievement-relevant information within the virtual student exams, six students' written answers were produced, which either were scientifically correct or expressed a specific misconception (i.e., anthropomorphic or teleological). The students' written answers are available in English as Additional material 1. The misconceptions articulated in the students' written answers were based on explanations from real students (Baalmann et al. 2004) and were modified for the SI. For reasons of homogeneity, all students' written answers were based on the Toulmin Argument Pattern (Toulmin 2003) and were tailored to the same length (115 words). The qualities of the written answers (i.e., anthropomorphic, teleological, or scientifically correct) were evaluated by three independent experts and resulted in an inter-rater agreement of 94%. According to AERA et al. (2014), this result can be interpreted as an indicator of content validity and gives an indication of the fit between the test items (i.e., the students' written answers) and the theoretical construct (i.e., the knowledge about the core ideas and principles of evolution: facet of CK; knowledge about student understanding: facet of PCK). Thus, the students' written answers clearly expressed a specific category of misconception or a scientifically correct explanation, and we were able to assume that the students' written answers could be applied as a measure of the trainee teachers' diagnostic knowledge about the core ideas and principles of evolution (facet of CK) and knowledge about student understanding on evolution (facet of PCK). To investigate a possible judgment bias (knowledge about judgment accuracy, facet of PK; see above), students' performance on a multiple-choice test on evolution (bad or good

performance) was integrated into the virtual student exams as achievement-irrelevant information (see Table 4.1). This performance had already been completed by the students and only needed to be noticed by the trainee teachers. The multiple-choice test on evolution was a separate task in the virtual student exams, so student performance shown in it should actually not affect the scoring of the further task (i.e., students' written answers on the natural selection of the peppered moth). The diagnostic activities on the facets of knowledge about the core ideas and principles of evolution (facet of CK), knowledge about student understanding of evolution (facet of PCK), and knowledge about judgment accuracy (facet of PK) were related to procedural knowledge according to psychological approaches (e.g., Fenstermacher 1994), because diagnostic knowledge had to be applied in an explicit situation (i.e., assessing students' exams; Förtsch et al. 2018; Kaiser et al. 2015).

Tabelle 4.1 Scheme of the variation of qualities of students' written answers and of the students' performance in the multiple-choice test in the SI.

Performance in the Multiple-Choice Test	Quality of Written Answers		
	<i>Teleological</i>	<i>Anthropomorphic</i>	<i>Scientifically correct</i>
Multiple-Choice: Good	<i>2 exams per variation</i> <i>(i.e., a total of 12 examinations)</i>		
Multiple-Choice: Bad			

4.4.3 Assessment of Declarative and Procedural Knowledge with the SI

In order to get first indications about the interrelationship between the declarative and procedural knowledge of trainee teachers about the facets of professional knowledge on evolution, we additionally integrated a short questionnaire with evolution-specific questions into the SI. This questionnaire on evolution consisted of 12 items, of which seven items were assigned to the CK domain and five items to the PCK domain. The items were taken from questionnaires previously used in other studies (KiL: Kleickmann et al. 2014; ProwiE: Großschedl et al. 2015). Two translated items are available in the Additional material 2. In the domain of CK, questions were asked about speciation, adaptation, and different evolutionary theories (Darwin, Lamarck). The PCK items focused on the reasons for misconceptions among students and the diagnosis of specific categories of misconceptions (e.g., anthropomorphic and teleological misconceptions). The knowledge required to answer the questions was classified as declarative knowledge because it is considered to be part of expert knowledge, which is explicit and learnable in academic

discourse (Baumert and Kunter 2013). The psychometric characteristics of the questionnaire on evolution (CK, PCK) were satisfactory (Cronbach's $\alpha = .63$). The aim was not to analyze the separate knowledge facets of CK and PCK, but to reveal the declarative knowledge of the trainee teachers. Accordingly, we considered the entire scale of items and we operationalized the results as the declarative professional knowledge of trainee teachers on evolution integrating the respective CK and the PCK. The teachers' diagnoses of the students' written answers in the SI were operationalized as the procedural knowledge of the trainee teachers because, here, knowledge had to be applied in an explicit action-oriented situation (i.e., assessing students' written answers; Förtsch et al. 2018; Kaiser et al. 2015; Schneider and Stern 2010).

4.4.4 Procedure

The trainee teachers needed approximately 60 minutes to complete the SI. At the beginning of the survey, the trainee teachers received short instructions on the SI. Each trainee teacher received six randomly selected virtual student exams. Each exam included the student's previous performance in the multiple-choice test about evolution (achievement-irrelevant information) and a written answer by the student on the natural selection of the industrial melanism of the peppered moth (achievement-relevant information), which included a misconception or a scientifically correct way of thinking. For an overview of a virtual student exam in the SI, see the Additional material 3. The performance in the previous multiple-choice test had already been assessed (good or bad multiple-choice performance) and was given to the trainee teachers without further information. Here, the trainee teachers only had to add up the points in the multiple-choice test to receive the final result for each student. If the students achieved 12 out of 20 points, this indicated a bad multiple-choice performance, whereas 19 out of 20 points indicated a good multiple-choice performance. The main task of the trainee teachers was to evaluate the students' written answers, whereby two written answers were scientifically correct and four written answers contained a misconception (i.e., two anthropomorphic and two teleological). A further task of the trainee teachers was to make sure that the students' written answers were assessed independent of the previous performance (i.e., multiple-choice test). The evaluation of the students' written answers included a scoring (i.e., between 0 and 20 points) and the diagnosis of the quality out of a list including five options (qualities used: anthropomorphic, teleological, scientifically correct; distractors: essential, religious). After the six exams of the students had been assessed by the trainee teachers, the trainee teachers completed a questionnaire in which their knowledge about the core ideas and principles of evolution (facet of CK; seven items) and their knowledge about student understanding (facet of PCK; five items) with regard to evolution was measured. Finally, the demographic information of the trainee teachers (e.g., age, gender, course of study, teaching experience) was recorded before the study was completed.

4.4.5 Analyses

Within the student exams, the achievement-relevant information (i.e., quality of students' written answers) was randomly combined with the achievement-irrelevant information (i.e., previous performance in the multiple-choice test), so that each quality type of written answer was combined with each performance in the multiple-choice test. The design was a fully crossed 2x3 design. This resulted in two independent variables (IV): (1) the quality (i.e., anthropomorphic, teleological, or scientifically correct) and (2) performance in the multiple-choice test (i.e., good or bad performance). The dependent variables (DV) were: (1) the trainee teachers' content-specific knowledge in the different facets of professional knowledge (i.e., diagnosis of the scientific correctness, facet of CK; diagnosis of specific misconceptions, facet of PCK) and (2) the trainee teachers' content-independent facet of professional knowledge (i.e., assessment of achievement-relevant information, facet of PK). A total of 162 student exams were included in the analysis (i.e., 27 trainee teachers who each had to assess six student exams). Relative frequencies in the diagnosis of scientific correctness and of the misconception category provided information about the biology-specific facets of professional knowledge within the CK (refers to RQ 1) and PCK domains (refers to RQ 2). The analysis of variance provided evidence of whether students' previous performance in a multiple-choice test on evolution (achievement-irrelevant information) influenced the assessment of the students' written answers (achievement-relevant information) within the virtual student exams (refers to RQ 3). Correlation analyses were used to obtain first indications about the interrelationship between the declarative knowledge (the trainee teachers' performance in the questionnaire) and the procedural knowledge (the trainee teachers' diagnosis of the students' written answers) of trainee teachers (refers to RQ 4).

4.5 Results

Research Question 1: The CK of Trainee Teachers

The trainee teachers were able to distinguish between scientifically correct (90.8% diagnosis rate) and scientifically incorrect (91.7% diagnosis rate) students' written answers. Accordingly, the analyses of variance revealed a significant main effect of the *quality of students' written answers* ($F[2,25] = 78.65, p < .001, \eta^2 = .863$; big effect), which means that scientifically correct written answers were scored higher than written answers with a misconception.

Research Question 2: The PCK of Trainee Teachers

Within the scientifically incorrect students' written answers, 61.1% of the anthropomorphic misconceptions and 27.8% of the teleological misconceptions were correctly diagnosed.

STUDIE 1

Overall, the trainee teachers correctly diagnosed 44.4% of the misconceptions into the respective misconception category. The results thus indicate a significant difference between the diagnosis of scientifically correct written answers and written answers that expressed a specific misconception category ($\chi^2[1] = 43.29, p < .001, \phi = .517$).

Research Question 3: The PK of Trainee Teachers

The study indicates a significant main effect of the previous performance in the multiple-choice test ($F[1,26] = 5.94, p < .022, \eta^2 = .186$; small effect; see Figure 4.2) on the judgment of the students' written answers. This means that scientifically correct students' written answers that were combined with a good multiple-choice performance received a higher score ($M = 18.63, SD = 1.80, \text{Min} = 15.00, \text{Max} = 20.00$) than students' written answers of the same quality that were combined with a bad multiple-choice performance ($M = 18.00, SD = 2.40, \text{Min} = 12.00, \text{Max} = 20.00$). Similar results were obtained for students' written answers with teleological misconceptions (good multiple-choice performance: $M = 9.78, SD = 4.73, \text{Min} = .00, \text{Max} = 19.00$; bad multiple-choice performance: $M = 7.81, SD = 3.25, \text{Min} = 2.00, \text{Max} = 14.00$). Students' written answers with anthropomorphic misconceptions received, on average, the same scores, regardless of the multiple-choice performance (good multiple-choice performance: $M = 7.19, SD = 4.28, \text{Min} = 2.00, \text{Max} = 16.00$; bad multiple-choice performance: $M = 7.19, SD = 3.94, \text{Min} = .00, \text{Max} = 16.00$). No interaction effect was found between previous performance in the multiple-choice test and quality of the written answers ($F[2,25] = .234, p < .05, ns$).

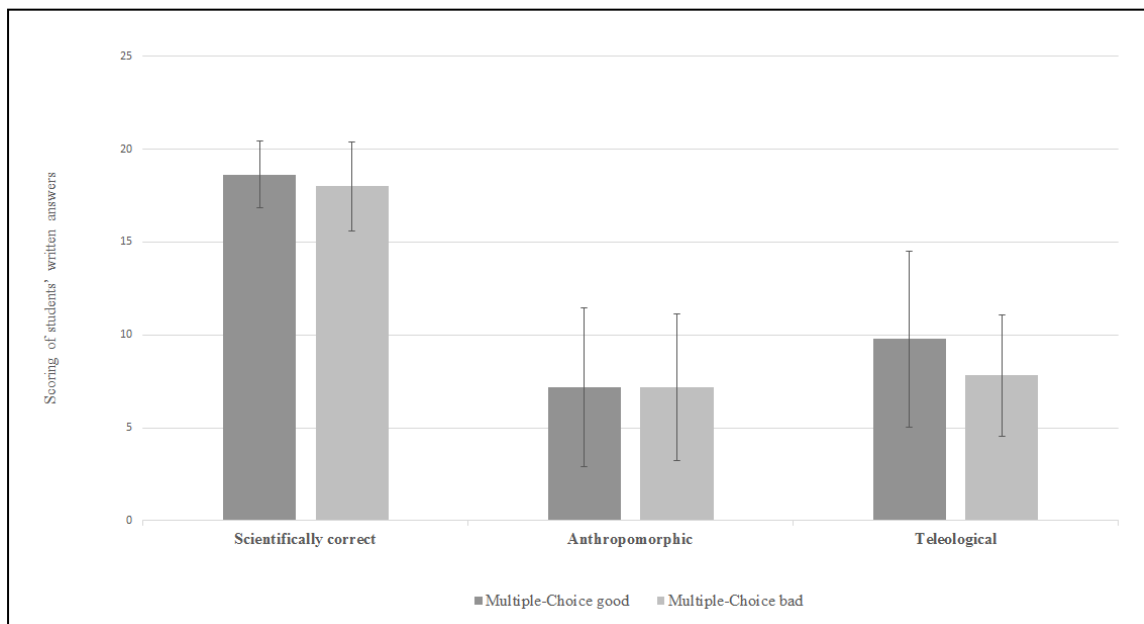


Abbildung 4.2 Main effect of the previous performance in the multiple-choice test.

Research Question 4: The PK of Trainee Teachers

The correlation between the trainee teachers' performance in the questionnaire and the diagnosis given in the CK domain (diagnosis of scientifically correct and scientifically incorrect students' written answers) revealed a strong effect ($r = .631, p < .001$). A moderate effect ($r = .499, p < .008$) was observed between the performance in the questionnaire and the diagnosis of the specific misconception category (PCK).

4.6 Discussion

The present study examined different facets of the professional knowledge of trainee teachers, specifically their knowledge of the core ideas and principles of evolution (facet of CK), their knowledge about student understanding (facet of PCK), and their knowledge about judgment accuracy (facet of PK). The trainee teachers used a digital instrument – the SI – to assess virtual student exams on evolution in biology. We used students' biology exams to combine achievement-relevant information (i.e., students' written answers with different quality levels: scientifically correct, anthropomorphic, teleological) with achievement-irrelevant information (previous performance in a multiple-choice test on evolution: good or bad performance). This experimental setting allowed us to investigate the trainee teachers' diagnosis in the facet of CK (i.e., diagnosis of scientific correctness), the facet of PCK (i.e., diagnosis of a specific misconception), and the facet of PK (i.e., diagnosis of achievement-relevant information). Based on a questionnaire, we were able to obtain first indications about the interrelationship between declarative and procedural knowledge.

The results show that trainee teachers were able to distinguish between scientifically correct and scientifically incorrect written answers in student exams. The diagnosis rate of both quality levels was over 90%. The assessment accuracy of the scientific quality level provided an indication of the trainee teachers' CK, which includes the knowledge facet of diagnosing the core ideas and principles of evolution (Großschedl et al. 2015). Research has shown that CK alone is not sufficient to ensure adaptive learning and the learning success of students (e.g., Abell 2007; Baumert et al. 2010; Förtsch et al. 2018). Diagnosing specific misconceptions about evolution requires knowledge about student understanding (facet of PCK) from trainee teachers and can be helpful in providing insights into the origin of unreflected and scientifically incorrect student answers. Trainee teachers diagnosed 44.4% of the specific misconception categories (i.e., anthropomorphic or teleological) in the student exams. Overall, this low diagnosis rate of misconceptions reveals a lack of the important facet of PCK, the knowledge about student understanding, among trainee teachers. These findings are in line with previous research that even college students have

STUDIE 1

problems in understanding evolutionary concepts (Alters and Nelson 2002). The diagnosis of teleological misconceptions (diagnosis rate: 27.8%) was significantly less frequent than that of anthropomorphic misconceptions (diagnosis rate: 61.1%). Moreover, when a student's written answer with a teleological misconception was diagnosed, trainee teachers rated this student's written answer better (i.e., a higher overall score) than a student's written answer with an anthropomorphic misconception. A study by Zohar et al. (1998) revealed that most high school students reason about biological phenomena with a mixture of teleological and causal reasoning. This reasoning is based on the tendency of humans to explain natural phenomena causally (e.g., Olander 2012). Consequently, processes that are explained on the basis of a goal- and purpose-oriented explanation are more likely to be accepted and may also be diagnosed less frequently as being scientifically incorrect by trainee teachers (e.g., Gregory 2009; Gresch and Martens 2019). The acceptance of teleological misconceptions manifests itself not only in the use of everyday language but also in unconscious actions in the classroom. This leads to teleology being an obstacle to understanding and explaining evolutionary processes (e.g., Evans et al. 2012; Kampourakis and Zogza 2008; Kelemen 2012; Sinatra et al. 2008). In contrast, trainee teachers scored students' written answers with anthropomorphic misconceptions significantly lower than students' written answers with teleological misconceptions. These findings indicate that trainee teachers consider anthropomorphic misconceptions to be more scientifically incorrect. The results of our study also confirm previous research that revealed that teachers often use teleological and anthropomorphic misconceptions to explain evolutionary processes (e.g., Kallery and Psillos 2004). The enormity of the challenge facing biologists and educators to diagnose the widespread misconception of natural selection is matched only by the importance of this task (Gregory 2009).

Independent of the articulated misconceptions, students' previous performance (i.e., performance in the multiple-choice test) influenced teachers' assessments of the students' written answers and revealed a lack of judgment accuracy (i.e., a halo effect). Thus, the students' written answers were rated higher if the corresponding performance in the previous multiple-choice test was good, indicating a judgment error (based on their PK; e.g., Südkamp et al. 2012). Achievement-irrelevant information was not directly related to performance but was often taken into account by teachers when assessing student performance, as shown in some previous studies (e.g., Ready and Wright 2011; Ritts et al. 1992; Schrader and Helmke 2001). This judgment error has also been reported in other studies that used the SI with pre-service, trainee, and in-service teachers in English and mathematics (Jansen et al. 2019; Kaiser et al. 2015). To the best of our knowledge, our results are the first to provide an indication that the halo effect also occurs in trainee teachers of biology. These results are particularly relevant because they show that, even after completing university education in biology, trainee teachers with bachelor's or

master's degrees are influenced by achievement-irrelevant information that distorts their judgments.

The present study provides first indications of the extent to which declarative knowledge (i.e., performance in a questionnaire on evolution) and procedural knowledge (i.e., diagnosis of scientific correctness or of specific misconceptions in students' written answers on natural selection) are interrelated among trainee teachers. Our results indicate that knowledge on evolution that is gained in the academic career can be transferred into specific action-oriented situations (e.g., assessing students' written answers on natural selection) and can generate procedural knowledge (Blömeke et al. 2010). Especially in the context of the theory of evolution, it is important to focus on and promote the proceduralization of declarative knowledge in university teacher education. In the future, further digital systems could offer new opportunities to investigate and close the existing theory-practice gap (e.g., Grossman and McDonald 2008). This may help to prepare pre-service teachers for the complexity of future classroom situations and to foster procedural knowledge on evolution.

In summary, natural selection is a key mechanism of modern evolutionary theory, which – in turn – is the connecting theme of all biology topics. Without a sophisticated understanding of this process and its consequences, it is simply impossible to even remotely understand the diversity of life. Thus, professional knowledge on evolution that is conveyed in university education must be focused in order to build an adequate knowledge base among pre-service teachers and, ultimately, to support students' learning of the concept of evolution.

4.6.1 Limitations

The SI we used presented students' written answers that dealt with the natural selection of the peppered moth and that contained various misconceptions and scientifically correct ways of thinking. Research shows that, depending on the biological organism (bacteria, plant, animal, or human), different misconceptions have a different probability of occurrence. As we used only texts on a zoological organism in the SI, we were not able to capture these dependencies.

Numerous misconceptions that can hinder the teaching and learning of evolution have been described in scientific research (e.g., Bishop and Anderson 1990; Gregory 2009; Settlage 1994; Nehm and Schonfeld 2008). In our study we applied just two common misconceptions (i.e., anthropomorphic and teleological misconceptions) that teachers often use to explain the process of natural selection.

The SI was processed by trainee teachers who had completed their university education, which mainly focuses on the development of declarative CK, PCK, and PK, with

the exception of short phases of practice in which procedural knowledge can be applied. Independent of the short practical training in the first phase of the teacher education program, the in-service teacher training program gives the trainee teachers the opportunity to gain in-depth teaching experience for the first time. Accordingly, it can be assumed that trainee teachers had not yet had much time to apply and train the declarative knowledge domains of PCK and PK in action-related situations (Blömeke et al. 2008). A sample of experienced teachers who have several years of practical experience could therefore lead to different results, which could indicate, for example, higher knowledge of student understanding (facet of PCK; Clermont et al. 1994; Grossmann 1990; Lederman et al. 1994; Schmelzing et al. 2013; van Driel et al. 2002) and generally more accurate judgments (facet of PK; Blömeke et al. 2015; Edelenbos and Kubanek-German 2004, Jansen et al. 2019, 2021).

4.6.2 Implications for Further Research and Practical Implications

The professional knowledge and the related diagnostic activities of pre-service and trainee teachers remain key aspects in biology education research. Education in schools as well as the university education of future teachers must place great importance on the theory of evolution in order to increase the awareness of the numerous misconceptions, as our results indicate that teachers face problems in dealing with their own misconceptions, which are actually similar to those of high school students (Demastes et al. 1995).

Misconceptions have been intensively researched (Gregory 2009). However, few studies have conceptualized which follow-up actions (e.g., cognitive conflict) are required to confront misconceptions (e.g., Demastes et al. 1995). Further research should focus on what action should be taken after the diagnosis of misconceptions. Here, the second facet of PCK, which contains knowledge about instructional strategies, is relevant. These instructional strategies can, depending on the diagnosed misconception, initiate and support the learning of the scientific concept (e.g., Ziadie and Andrews 2018). To counter students' misconceptions about evolution, the conceptual change theory, for example, is a suitable approach (Demastes et al. 1995). Here, conceptual change is described as the learning process from pre-instructional conceptions to the acquisition of scientific concepts (Heitz et al. 2010).

Initial analyses indicate a correlation between the declarative and procedural knowledge of trainee teachers. However, the extent to which declarative knowledge (i.e., performance in a questionnaire) is related to procedural knowledge (i.e., action-oriented knowledge) remains unclear. Therefore, further analyses should be conducted with more extensive questionnaires (i.e., with additional items) in order to gain a deeper insight into the interrelationships between the different types of knowledge. This could be achieved by using digital instruments, such as simulated classroom environments. The results obtained

from such analyses might help to reveal deficits in university teacher education and to indicate which specific contexts need to be targeted to bring declarative knowledge into action-oriented situations and to thereby generate procedural knowledge.

4.7 References

- Abell SK. Research on science teacher knowledge. In: Abell SK, editor. *Handbook of research on science education*. NJ: Lawrence Erlbaum Associates; 2007. p. 1105–1149.
- AERA, APA, NCME. *Standards for educational and psychological testing*. Washington: American Educational Research Association. 2014.
- Allmon WD. Why don't people think evolution is true? Implications for teaching, in and out of the classroom. *Evol Educ Outreach*. 2011;4(4):648–65.
- Alters BJ, Nelson CE. Perspective: Teaching evolution in higher education. *Evol Int J Org Evol*. 2002;56(10):1891–901.
- Anderson DL, Fisher KM, Norman GJ. Development and evaluation of the conceptual inventory of natural selection. *J Res Sci Teach*. 2002;39(10):952–78.
- Andersson B, Wallin A. On developing content-oriented theories taking biological evolution as an example. *Int J Sci Educ*. 2006;28(6):673–95.
- Andrews TM, Kalinowski ST, Leonard MJ. "Are humans evolving?" A classroom discussion to change student misconceptions regarding natural selection. *Evol Educ Outreach*. 2011;4(3):456–66.
- Athanasiou K, Papadopoulou P. Conceptual ecology of the evolution acceptance among Greek education students: Knowledge, religious practices and social influences. *Int J Sci Educ*. 2012;34(6):903–24.
- Baalmann W, Frerichs V, Weitzel H, Gropengießer H, Kattmann U. Schülervorstellungen zu Prozessen der Anpassung–Ergebnisse einer Interviewstudie im Rahmen der Didaktischen Rekonstruktion (Students' conceptions of processes of adaptation results of an interview study in the context of didactic reconstruction). *ZfDN*. 2004;10(1):7–28.
- Barnes ME, Evans EM, Hazel A, Brownell SE, Nesse RM. Teleological reasoning, not acceptance of evolution, impacts students' ability to learn selection. *Evol: Edu and Outreach*. 2017;10(7):1–12.
- Basel N, Harms U, Prechtel H. Analysis of students' arguments on evolutionary theory. *J Biol Educ*. 2013;47(4):192–9.

STUDIE 1

- Basel N, Harms U, Precht H, Weiss T, Rothgangel M. Students' arguments on the science and religion issue: the example of evolutionary theory and genesis. *J Biol Educ.* 2014;48(4):179–87.
- Baum DA, DeWitt-Smith S, Donovan SSS. The tree-thinking challenge. *Sci.* 2005;310(5750):979–80.
- Baumert J, Kunter M. The COACTIV model of teachers' professional competence. In *Cognitive activation in the mathematics classroom and professional competence of teachers*. Boston: Springer; 2013. p. 25–48
- Baumert J, Kunter M, Blu W, Brunner M, Voss T, Jordan A, Klusmann U, Kauss S, Neubrand M, Tsai YM. Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *Amer Educ Res J.* 2010;47:133–80.
- Beardsley PM. Middle school student learning in evolution: Are current standards achievable? *Amer Biol Tea.* 2004;66(9):604–13.
- Berkman MB, Plutzer E. Defeating creationism in the courtroom, but not in the classroom. *Sci.* 2011;331(6016):404–5.
- Berti AE, Toneatti L, Rosati V. Children's conceptions about the origin of species: A study of Italian children's conceptions with and without instruction. *J Lear Sci.* 2010;19(4):506–38.
- Billingsley B, Brock R, Taber KS, Riga F. How students view the boundaries between their science and religious education concerning the origins of life and the universe. *Sci Educ.* 2016;100(3):459–82.
- Bishop BA, Anderson CW. Student conceptions of natural selection and its role in evolution. *J Res Sci Teach.* 1990;27(5):415–27.
- Blömeke S, Felbrich A, Müller C, Kaiser G, Lehmann R. Effectiveness of teacher education. State of research, measurement issues and consequences for future studies. *Int J Math Educ.* 2008;40(5):719–34.
- Blömeke S, Kaiser G, Lehmann R. TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich (TEDS-M 2008. Professional competence and learning opportunities of future primary school teachers in international comparison). Münster: Waxmann; 2010.
- Blömeke S., Gustafsson JE, Shavelson RJ. Beyond dichotomies. *Zeitschrift für Psychologie.* 2015;223:3–13.
- Borman WC. Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings. *J App Psy.* 1975;60(5):556–60.

- Bruckermann T, Fiedler D, Harms U. (2020). Identifying precursory concepts in evolution during early childhood—a systematic literature review. *Stud Sci Educ.*2020;1–43.
- Brumby MN. Misconceptions about the concept of natural selection by medical biology students. *Sci Educ.* 1984;68(4):493–503.
- Brunner M, Anders Y, Hachfeld A, Krauss S. Diagnostische Fähigkeiten von Mathematiklehrkräften (Diagnostic abilities of mathematics teachers). In: Kunter M, Baumert J., Blum W, Klusmann U, Krauss S, Neubrand M, editors. *Professionelle Kompetenz von Lehrkräften - Ergebnisse des Forschungsprogramms COACTIV* (Professional competence of teachers - results of the COACTIV research program). Münster: Waxmann; 2011. p. 215–34.
- Byrne J, Grace M, Hanley P. Children's anthropomorphic and anthropocentric ideas about micro-organisms: Educational research. *J Biol Educ.* 2009;44(1):37–43.
- Catley KM, Phillips BC, Novick LR. Snakes and eels and dogs! Oh, my! Evaluating high school students' tree-thinking skills: An entry point to understanding evolution. *Res Sci Educ.* 2013;43:2327–48.
- Clermont CP, Borko H, Krajcik JS. Comparative study of the pedagogical content knowledge of experienced and novice chemical demonstrators. *J Res Sci Teach.* 1994;31:419–41.
- Dagher ZR, BouJaoude S. Scientific views and religious beliefs of college students: The case of biological evolution. *J Res Sci Teach.* 1997;34(5):429–45.
- Demastes SS, Settlage J, Good R. Students' conceptions of natural selection and its role in evolution: Cases of replication and comparison. *J Res Sci Teach.* 1995;32(5):535–50.
- Deniz H, Donnelly LA, Yilmaz I. Exploring the factors related to acceptance of evolutionary theory among Turkish preservice biology teachers: Toward a more informative conceptual ecology for biological evolution. *J Res Sci Teach.* 2008;45:420–43.
- van Dijk EM, Kattmann U. Evolution im Unterricht: Eine Studie über fachdidaktisches Wissen von Lehrerinnen und Lehrern (Evolution in teaching: A study of the didactic knowledge of teachers). *ZfDN.* 2010;16:7–21.
- van Driel JH, De Jong O, Verloop N. The development of preservice chemistry teachers' pedagogical content knowledge. *Sci Edu.* 2002;86:572–590.
- Driver R. Changing conceptions. *J Res Educ.* 1988;161–98.
- Edelenbos P, Kubanek-German A. Teachers assessment: The concept of diagnostic competence. *Language Testing.* 2004;21(3):259–83.

STUDIE 1

- Endler JA. Natural selection in the wild. Monographs in population biology: Vol 21. Princeton University Press.
- Evans EM. The emergence of beliefs about the origins of species in school-age children. *Merr-Palm Quart.* 2000;46(2):221–254.
- Evans EM, Rosengren K, Lane JD, Price KLS. Encountering counterintuitive ideas: Constructing a developmental learning progression for evolution understanding. In: Rosengren K., Brem SK, Evans EM, Sinatra GM, editors. *Evolution challenges: Integrating research and practice in teaching and learning about evolution.* New York: Oxford University Press, 2012. p. 174–99.
- Fenstermacher GD. The knower and the known: The nature of knowledge in research on teaching. *Rev Res Educ.* 1994;20(1),3–56.
- Ferrari M, Chi MT. The nature of naive explanations of natural selection. *Int J Sci Educ.* 1998;20(10):1231–256.
- Fiedler D, Sbeglia GC, Nehm RH, Harms U. How strongly does statistical reasoning influence knowledge and acceptance of evolution? *J Res Sci Teach.* 2019;56(9):1183–206.
- Fiedler D, Tröbst S, Harms U. University students' conceptual knowledge of randomness and probability in the contexts of evolution and mathematics. *CBE Life Sci Educ.* 2017;16(2):1–16.
- Fiedler D, Tröbst S, Großschedl J, Harms U. EvoSketch: Simple simulations for learning random and probabilistic processes in evolution, and effects of instructional support on learners' conceptual knowledge. *Evol Educ Outreach.* 2018;11(15):1–17.
- Friedrichsen PJ, Abdell S, Pareja E, Brown P, Lankford D, Volkmann M. Does teaching experience matter? Examining biology teachers' prior knowledge for teaching in an alternative certification program. *J Res Sci Teach.* 2009;46(4):357–83.
- Furtak EM. Linking a learning progression for natural selection to teachers' enactment of formative assessment. *J Res Sci Teach.* 2012;49(9):1181–210.
- Förtsch C, Sommerhoff D, Fischer F, Fischer M, Girwidz R, Obersteiner A, Reiss K, Stürmer K, Siebeck M, Schmidmaier R, Seidel T, Ufer S, Wecker C, Neuhaus BJ. Systematizing Professional Knowledge of Medical Doctors and Teachers: Development of an Interdisciplinary Framework in the Context of Diagnostic Competences. *Educ Sci.* 2018;8(4):207–25.
- Garvin-Doxas K, Klymkowsky MW. Understanding randomness and its impact on student learning: lessons learned from building the Biology Concept Inventory (BCI). *CBE Life Sci Educ.* 2008;7(2):222–37.

- German National Academy of Sciences Leopoldina. Evolutionsbiologische Bildung in Schule und Universität (Evolutionary biology education in schools and universities). Halle, Germany; 2017.
- Godfrey-Smith P. Conditions for evolution by natural selection. *J Philos.* 2007;104(10):489–516.
- Gregory TR. Understanding evolutionary trees. *Evol Educ Outreach.* 2008;1(2),121– 37.
- Gregory TR. Understanding Natural selection. Essential concepts and common misconceptions. *Evol Educ Outreach.* 2009;2(2):156–75.
- Gresch H, Martens M. *Teleology as a tacit dimension of teaching and learning evolution: A sociological approach to classroom interaction in science education.* *J Res Sci Teach.* 2019;56(3):243–69.
- Grossmann PL. The making of a teacher: Teacher knowledge and teacher education. New York: Teacher College Press; 1990.
- Grossman PL, McDonald M. Back to the Future: Directions for Research in Teaching and Teacher Education. *Amer Educ Res J.* 2008;45(1):184–205.
- Großschedl J, Harms U, Kleickmann T, Glowinski I. Preservice biology teachers' professional knowledge: Structure and learning opportunities. *J Sci Teach Educ.* 2015;26(3):291–318.
- Großschedl J, Konnemann C, Basel N. Pre-service biology teachers' acceptance of evolutionary theory and their preference for its teaching. *Evol Educ Outreach.* 2014;7(18):1–16.
- Großschedl J, Seredszus F, Harms U. Angehende Biologielehrkräfte: evolutionsbezogenes Wissen und Akzeptanz der Evolutionstheorie (Prospective biology teachers: evolution-related knowledge and acceptance of the theory of evolution). *ZfDN.* 2018;24(1):51–70.
- Ha M, Baldwin BC, Nehm RH. The long-term impacts of short-term professional development: Science teachers and evolution. *Evol Edu Outreach.* 2015;1–23.
- Ha M, Haury DL, Nehm RH. Feeling of certainty: Uncovering a missing link between knowledge and acceptance of evolution. *J Res Sci Teach.* 2012;49(1):95–121.
- Harms U, Reiss MJ. *Evolution Education Re-considered: Understanding What Works.* Basel: Springer; 2019.
- Hashweh MZ. Teacher pedagogical constructions: A reconfiguration of pedagogical content knowledge. *Teachers and Teaching.* 2005;11:273–92.

STUDIE 1

- Heitz JG, Cheetham JA, Capes EM, Jeanne RL. Interactive evolution modules promote conceptual change. *Evol Educ Outreach*. 2010;3(3):436–42.
- Helmke A, Hosenfeld I, Schrader FW. Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften (Comparative studies as an instrument for improving the diagnostic competence of teachers). In: Arnold R, Grieser C, editors. *Schulleitung und Schulentwicklung*. Hohengehren: Schneider-Verlag; 2004. p. 119–44.
- Hill HC, Ball DL, Schilling SG. Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *J Res Math Educ*. 2008;39:372–400.
- Jansen T, Vögelin C, Machts N, Keller S, Möller J. Das Schülerinventar ASSET zur Beurteilung von Schülerarbeiten im Fach Englisch: Drei experimentelle Studien zu Effekten der Textqualität und der Schülernamen (The student inventory ASSET for evaluating student work in English: Three experimental studies on effects of text quality and student names). *Psychologie in Erziehung und Unterricht*. 2019;66(4),303–15.
- Jansen T, Vögelin C, Machts N, Keller S, Köller O, Möller J. Judgment accuracy in experienced versus student teachers: Assessing essays in English as a foreign language. *Teach Teacher Educ*. 2021;97:1–11.
- Kaiser J, Möller J, Helm F, Kunter M. Das Schülerinventar: Welche Schülermerkmale die Leistungsurteile von Lehrkräften beeinflussen (The Student Inventory: What Student Characteristics Influence Teacher Performance Judgments). *Zeitschrift für Erziehungswissenschaft*. 2015;18(2):279–302.
- Kaiser J, Südkamp A, Möller J. The effects of student characteristics on teachers' judgment accuracy: Disentangling ethnicity, minority status, and achievement. *J Educ Psychol*. 2017;109(6):871–88.
- Kallery M, Psillos D. Anthropomorphism and animism in early years science: Why teachers use them, how they conceptualise them and what are their views on their use. *Res Sci Educ*. 2014;34(3):291–311.
- Kampourakis K, Zogza V. Students' intuitive explanations of the causes of homologies and adaptations. *Sci Educ*. 2008;17(1):27–47.
- Kampourakis K, Nehm RH. History and philosophy of science and the teaching of evolution: Students' conceptions and explanations. *International Handbook of Research in History, Philosophy and Science Teaching*. Dordrecht: Springer; 2014. p. 377–99.

- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedem.12000>
- Kelemen D. Teleological minds: How natural intuitions about agency and purpose influence learning about evolution. In: Rosengren K, Evans EM. Editors. *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. England: Oxford University Press; 2012; 66–92.
- King CJH. An analysis of misconceptions in science textbooks: Earth science in England and Wales. *Intl J Sci Educ*. 2010;32(5):565–601.
- Kleickmann T, Großschedl J, Harms U, Heinze A, Herzog S, Hohenstein F, Köller O, Kröger J, Lindmeier A, Loch C, Mahler D, Möller J, Neumann K, Parchmann I, Steffensky M, Taskin V, Zimmermann F. Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer – Testentwicklung im Rahmen des Projekts KiL (Professional knowledge of pre-service teachers in mathematics and science - Test development within the framework of the KiL project). *Unterrichtswissenschaft*. 2014;42:280–8.
- Kunter M, Klusmann U, Baumert J, Richter D, Voss T, Hachfeld A. Professional competence of teachers: Effects on instructional quality and student development. *J Educ Psychol*. 2013;105(3):805–20.
- König J, Blömeke S, Klein P, Suhl U, Busse A, Kaiser G. Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teach Teacher Educ*. 2014;38:76–88.
- Lederman NG, Gess-Newsome J, Latz MS. The nature and development of preservice science teachers' conceptions of subject matter and pedagogy. *J Res Sci Teach*. 1994;31:129–46.
- Lee E, Luft JA. Experienced secondary science teachers' representation of pedagogical content knowledge. *Int J Sci Educ*. 2008;30:1343–63.
- Magnusson S, Krajcik J, Borko H. Nature, sources, and development of pedagogical content knowledge for science teaching. In: *Examining pedagogical content knowledge*. Dordrecht: Springer; 1999. p. 95–132.
- Malouff JM, Thorsteinsson EB. Bias in grading: A meta-analysis of experimental research findings. *Australian J Educ*, 2016;60(3),245–56.
- Mayr E. *The growth of biological thought: Diversity, evolution and inheritance*. Cambridge: Harvard University Press; 1982.

STUDIE 1

- Mahler, D., Großschedl, J., & Harms, U. Using doubly latent multilevel analysis to elucidate relationships between science teachers' professional knowledge and students' performance. *Int J Sci Educ.* 2017;39(2):213–37.
- McVaugh NK, Birchfield J, Lucero MM, Petrosino AJ. Evolution education: Seeing the forest for the trees and focusing our efforts on the teaching of evolution. *Evol Educ Outreach.* 2011;4(2):286–92.
- Moharrer K, Ha M, Nehm R. EvoGrader: an online formative assessment tool for automatically evaluating written evolutionary explanations. *Evol Edu Outreach.* 2014;7(1):15.
- Murphy KR, Reynolds DH. Does true halo affect observed halo? *J Appl Psychol.* 1988;73(2):235–8.
- Nadelson LS, Sinatra GM. Educational professionals' knowledge and acceptance of evolution. *Evol Psychol.* 2009;7(4):490–516.
- National Research Council. Thinking evolutionarily: Evolution education across the life sciences: Summary of a convocation. Washington, DC: The National Academies Press; 2012.
- Nehm RH, Poole TM, Lyford ME, Hoskins SG, Carruth L, Ewers BE, Colberg PJ. Does the segregation of evolution in biology textbooks and introductory courses reinforce students' faulty mental models of biology and evolution? *Evol Educ Outreach.* 2009;2(3):527–32.
- Nehm RH, Rector MA, Ha M. "Force-Talk" in evolutionary explanation: Metaphors and misconceptions. *Evol Educ Outreach.* 2010;3:605–13.
- Nehm RH, Reilly L. Biology majors' knowledge and misconceptions of natural selection. *BioScience.* 2007;57(3):263–72.
- Nehm RH, Schonfeld IS. Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *J Res Sci Teach.* 2008;45(10):1131–60.
- Nehm RH, Schonfeld IS. Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools? *J Sci Teach Educ.* 2007;18(5):699–723.
- Neubrand C. Untersuchung von Aspekten erfolgreichen Lehrens und Lernens mit Fokus auf Beispielaufgaben und vorwissensangepasste Impulse (Investigation of aspects of successful teaching and learning with focus on example tasks and pre-knowledge-adapted impulses) [Unpublished doctoral dissertation]. Kiel: Leibniz Institute for Science and Mathematics Education; 2017.

- Neumann K, Härtig H, Harms U, Parchmann I. Science Teacher Preparation in Germany: An International Comparison of What Works Best. In: Pederson J, Isozaki T, Hirano T, editors. *Science Teacher Education: A Multi Country Comparison*. Charlotte, NC: Information Age Publishing Inc; 2017. p. 29–52.
- NGSS Lead States [NGSS] (2013). *Next generation science standards: for states, by states*. Washington, DC: The National Academies Press.
- Olander C. Teaching biological evolution-internal and external evaluation of learning outcomes. *Nord Stud Sci Educ*. 2012;5(2):171–84.
- Opfer JE, Nehm RH, Ha M. Cognitive foundations for science assessment design: knowing what students know about evolution. *J Res Sci Teach*. 2012;49(6):744–77.
- Oudman S, van de Pol J, Bakker A, Moerbeek M, van Gog T. Effects of different cue types on the accuracy of primary school teachers' judgments of students' mathematical understanding. *Teaching Teacher Educ*. 2018;76:214–26.
- Park S, Oliver JS. Revisiting the conceptualization of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Res Sci Educ*. 2008;38:261–84.
- Pazza R, Penteado PR, Kavalco KF. Misconceptions about evolution in Brazilian freshmen students. *Evol Educ Outreach*. 2010;3(1):107–13.
- Phillips BC, Novick LR, Catley KM, Funk DJ. Teaching tree thinking to college students: It's not as easy as you think. *Evol Educ Outreach*. 2012;5(4):595–602.
- Ready DD, Wright DL. Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *Ameri Educ Res J*. 2011;48(2):335–60.
- Rector MA, Nehm RH, Pearl D. Learning the language of evolution: lexical ambiguity and word meaning in student explanations. *Res Sci Educ*. 2013;43(3):1107–33.
- Renkl A. Träges Wissen: Wenn Erlerntes nicht genutzt wird (Tacit knowledge: If learned knowledge is not used.) *Psychologische Rundschau*. 1996;47:78–92.
- Rissler LJ, Duncan SI, Caruso NM. The relative importance of religion and education on university students' views of evolution in the Deep South and state science standards across the United States. *Evol Educ Outreach*. 2014;7(1):7–24.
- Ritts V, Patterson ML, Tubbs ME. Expectations, impressions, and judgments of physically attractive students: A review. *Rev Educ Res*. 1992;62(4):413–26.
- Schmelzing S, van Driel JH, Jüttner M, Brandenbusch S, Sandmann A, Neuhaus BJ. Development, evaluation, and validation of a paper-and-pencil test for measuring

STUDIE 1

- two components of biology teachers' pedagogical content knowledge concerning the 'cardiovascular system'. *Int J Sci Math Edu.* 2013;11:1369–390.
- Schneider M, Stern E. The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Develop Psychol.* 2010;46(1):178–92.
- Schrader FW. Diagnostische Kompetenz von Eltern und Lehrern (Diagnostic competence of parents and teachers). In: Rost DH, editor. *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz; 2006. p. 95–100.
- Schrader FW, Helmke A. Alltägliche Leistungsbeurteilung durch Lehrer (Day-to-day performance evaluation by teachers). *Leistungsmessungen in Schulen.* 2001;2:45–58.
- Settlage J. Conceptions of natural selection: A snapshot of the sense-making process. *J Res Sci Teach.* 1994;31(5):449–57.
- Shtulman A. Qualitative differences between naïve and scientific theories of evolution. *Cog Psychol.* 2006;52(2):170–94.
- Shulman LS. Those who understand: Knowledge growth in teaching. *Educ Res.* 1986;15(2):4–14.
- Shulman LS. Knowledge and teaching: Foundations of the new reform. *Harvard Educ Rev.* 1987;57(1):1–22.
- Sickel AJ, Friedrichsen P. Examining the evolution education literature with a focus on teachers: major findings, goals for teacher preparation, and directions for future research. *Evol Educ Outreach.* 2013;6(1):6–23.
- Sinatra GM, Brem SK, Evans EM. Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evol Educ Outreach.* 2008;1(2):189–95.
- Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States in the Federal Republic of Germany [KMK]. *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss (Beschluss der Kultusministerkonferenz vom 16.12.2004) (Science standards for middle school biology)*. Munich, Germany: Luchterhand; 2005.
- Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States in the Federal Republic of Germany [KMK]. *Standards für die Lehrerbildung Bildungswissenschaften* Munich, Germany: Luchterhand; 2014.
- Stover SK, Mabry ML. Influences of Teleological and Lamarckian Thinking on Student Understanding of Natural Selection. *Bioscience.* 2007;33(1):11–8.

- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *J Educ Psychol.* 2012;104(3):743–62.
- Tamir P. Subject matter and related pedagogical knowledge in teacher education. *Teaching and Teacher Education.* 1988;4(2):99–110.
- Tamir P, Zohar A. Anthropomorphism and teleology in reasoning about biological phenomena. *Sci Educ.* 1991;75(1):57–67.
- Tepner O, Borowski A, Fischer HE, Jüttner M, Kirschner S, Leutner D. Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften (Model for the development of test items for the assessment of professional knowledge of science teachers). *Zeitschrift für Didaktik der Naturwissenschaft.* 2012;18:7–28.
- Tibell LA, Harms U. Biological principles and threshold concepts for understanding natural selection. *Sci Educ.* 2017;26(7):953–73.
- Toulmin SE. *The uses of argument.* Cambridge. UK: Cambridge University Press; 2003.
- Vögelin C, Jansen T, Keller S, Machts N, Möller J. The influence of lexical features on teacher judgments of ESL argumentative essays. *Assess Writing.* 2019;39:50–63.
- Wandersee JH, Good RG, Demastes SS. Forschung zum Unterricht über Evolution: Eine Bestandsaufnahme (Research on the teaching of evolution: an inventory). *ZfDN.* 1995;1(1):43–54.
- Yasri P, Mancy R. Understanding student approaches to learning evolution in the context of their perceptions of the relationship between science and religion. *Int J Sci Educ.* 2014;36(1):24–45.
- Yip DY. Identification of misconceptions in novice biology teachers and remedial strategies for improving biology learning. *Int J Sci Educ.* 1998;20(4):461–77.
- Zabel J, Gropengießer H. Learning progress in evolution theory: climbing a ladder or roaming a landscape? *J Biol Educ.* 2011;45(3):143–9.
- Ziadie MA, Andrews TC. Moving evolution education forward: a systematic analysis of literature to identify gaps in collective knowledge for teaching. *CBE Life Sci Educ.* 2018;17(1):1–10.
- Zohar, A., & Ginossar, S. Lifting the taboo regarding teleology and anthropomorphism in biology education—heretical suggestions. *Sci Educ.* 1998;82(6):679–97.

4.8 Supplemental Material

4.8.1 Overview About the Student Written Answers

Additional file 1: Task in the student virtual exam about the natural selection of the peppered moth and corresponding student virtual texts (i.e., students' written answers on the task) with different text qualities (i.e., anthropomorphic, teleological, scientifically correct) in the SI. The virtual students' written answers were given to the trainee teachers in German, which were translated into English for the additional material.

Task: Natural selection of the peppered moth

At the beginning of the 1950s, Bernard Kettlewell (1907-1979), a British geneticist and entomologist, engaged in his research with the different color-patterns of the peppered moth (*Biston betularia*). During this time, specific areas in England were exposed to massive industrial pollution, whereby as a result of an increased emission of exhaust fumes produced by factories (soot and sulfur oxide), the tree bark darkened. Kettlewell observed, surmising that there was an original population in these areas, that especially the darker type (*forma carbonaria*) asserted itself against the lighter type (*forma typica*) of the peppered moth. How can the dominance of the darker type (*forma carbonaria*) of the peppered moth in the areas with industrial pollution be explained according to evolutionary biology?

Student written answer 1 (anthropomorphic)

Before industrialization, white peppered moths had the possibility to hide on the light bark of the birch. Due to this camouflage, they were harder to find and were eaten less frequently by enemies. With the darkened tree bark, the lighter peppered moths had a selection disadvantage. They were less camouflaged and were therefore found and eaten more frequently by enemies. Over the course of time, the peppered moth realized that it couldn't survive with its light coloring and this led to the adaptation. Subsequently, body cells were activated, which dyed the peppered moth darker. Now they were able to camouflage themselves on the darker bark again. Eventually, more dark moths survived and could reproduce, while the lighter ones were frequently eaten by natural enemies.

Student written answer 2 (anthropomorphic)

Due to industrialization, the habitat of the peppered moths changed. The birch bark, which used to be white and which used to be the home of many light peppered moths, darkened, which mean that the moths were no longer camouflaged and they were thus found by enemies and eaten more frequently. Nature realized that the moths could not survive under these conditions and developed better peppered moths that were darker and thus better adapted to the environment. They were able to hide on the modified birch bark and were,

therefore, eaten less frequently. Like this, more and more dark peppered moths reproduced and spread throughout England. The dominance of the dark type can thus be explained through the will of nature, which arranged the adaptation of the peppered moths.

Student written answer 3 (teleological)

With the emergence of industrialization in England, the light tree barks of the birches darkened. Consequently, the natural habitat of the peppered moths, which had been light colored until then, did not provide sufficient camouflage anymore. The light individuals were easily found by their enemies and thus heavily depleted. Therefore, the coloring of the peppered moths changed to become darker. The darker coloring of the peppered moths now served the purpose of camouflaging them on the birch barks and thus protected them from potential enemies, such as birds. Because the peppered moth with the higher melanin content was preyed on less frequently because of its good camouflage, it became dominant in those areas that were severely affected by industrial pollution.

Student written answer 4 (teleological)

When the soot darkened the tree barks, the light peppered moths were easy prey for their natural enemies. The white coloring of the moth did not provide camouflage anymore and it no longer fulfilled its original purpose. Birds flying past could see the light-colored moths on the darker birch barks, which led to many of the moths being eaten. For this reason, the coloring of the peppered moths changed. The darker individuals could easily stay on the soot-covered tree barks without being seen by their enemies. Therefore, the darker coloring helped the peppered moths to camouflage themselves. This kind of adaptation eventually led to increased survival chances for the darker peppered moths and thus led to their dominance in this area.

Student written answer 5 (scientifically correct)

The soot-covered trees did not provide sufficient camouflage for the lighter individuals of the peppered moths anymore. For that reason, the lighter individuals were found and eaten by their natural enemies more frequently. The peppered moths that contained more melanin due to a variation in the original population had a selection advantage. Because of the darker coloring, which can be ascribed to mutation and a recombination of specific genes, it was harder for their natural enemies to find them on the modified birch bark and, thus, they were eaten less frequently. This mechanism led to a gradual decrease in the population of the lighter peppered moths, while the moths with more melanin took up the bigger part of the population. The gene pool of the peppered moth population changed in this area, which explains the dominance of the darker individuals.

Student written answer 6 (scientifically correct)

Because of the darkening of the white bark of the birch, the lighter types of the peppered moths were less camouflaged during the day. Consequently, the lighter peppered moths were found more frequently by their natural enemies on the trees, which were now polluted by soot, and were decimated. In contrast, there was also the darker type of the peppered moth, whose increased melanin content can be ascribed to mutation and a recombination of specific genes. The dark coloration of the birches led to better camouflage for the darker peppered moths, which could thus not be found as easily by their natural enemies. Therefore, they had a selection advantage as they were eaten less frequently. Consequently, their survival probability increased, and the gene pool shifted towards the peppered moths with an increased melanin content.

4.8.2 Examples of Items in the Questionnaire on Evolution

Additional file 2: The questionnaire on evolution consisted of 12 items, of which seven items were assigned to the CK domain and five items to the PCK domain. Two items are presented below; the first item captures knowledge in the CK domain (i.e., knowledge about core ideas and theories on evolution) and the second item captures knowledge in the PCK domain (i.e., knowledge about student understanding on evolution). Both items originate from a questionnaire for the assessment of evolution-specific professional knowledge in German (ProWiE: Großschedl et al. 2015). The following two items were translated into English.

Item from the CK domain:

Tick which of the following statements explains the effect of isolation mechanisms. (1 answer)

- ☐ Isolation mechanisms contribute to the development of different gene pools.
- ☐ Isolation mechanisms support coevolution.
- ☐ Isolation mechanisms impact the mutation rate.
- ☐ Isolation mechanisms increase non-heritable variability (modifications).
- ☐ Isolation mechanisms inhibit artificial selection.

Item from the PCK domain:

State the theory of evolution that emerges from the following student conception.

"If cheetahs need to run faster to get to their prey, and they want that very strongly, then they will be able to run faster."

4.8.3 Screenshots of Virtual Student Exams in the SI

Additional file 3: The screenshots represent the procedure for assessing virtual student exams in the Student Inventory (SI). The SI was programmed in German, and detailed explanations were added to each screenshot.

Screenshots 1 – 2: On the left side of the computer screen, the virtual student exam is displayed, whereas on the right side of the screen, the assessments can be entered by the trainee teacher. Here, the first task (i.e., the multiple-choice test on evolution) of the virtual student exam is shown. This multiple-choice task has already been answered by the students (i.e., tick marks); the trainee teachers only have to check the ticks made by the students in the multiple-choice task. The scoring scheme for all multiple-choice test questions is given to the trainee teacher on the right side of the screen (top half), so that they do not have to evaluate the answers themselves. In total, the multiple-choice test comprises six tasks and results in either a good (19 out of 20 points) or bad (12 out of 20 points) performance. Here, the virtual student achieved a good multiple-choice performance of 19 points (see Screenshot 2). This performance in the multiple-choice test is considered to be irrelevant to judgment, because it should have no influence on the assessment of the next task (see Screenshot 3).

Screenshots 3 – 4: The left side of the computer screen displays the task and the student's written answer on the natural selection of the peppered moth. On the right side, the trainee teachers can enter the scoring for the student's written answer (0 - 20 points) and must indicate whether the answer is scientifically correct or contains a misconception. This part of the virtual student exam is operationalized as achievement-relevant information, because, here, the trainee teachers have to apply their own knowledge in order to assess the students' written answers. The previously shown performance of the students in the multiple-choice test (achievement-irrelevant information) should have no effect on the assessment of the student's written answers (achievement-relevant information), because these were to be assessed separately from each other. Here, the student's written answer was scientifically correct, resulting in the teacher scoring the answer (see Screenshot 4) as good (19 points).

Within the virtual student exams, each quality of the students' written answers (scientifically correct, anthropomorphic, teleological; achievement-relevant information) was randomly combined with one of the two multiple-choice performances (good or bad performance; achievement-irrelevant information). Thus, each quality of written answer was combined with each performance on the multiple-choice test. Accordingly, each trainee teacher had to assess 6 virtual student exams in the SI.

Abbildung 4.3 Screenshot 1 of the SI

Abbildung 4.4 Screenshot 2 of the SI

7. Anfang der 50er Jahre beschäftigte sich Bernard Kettlewell (1907-1979), ein britischer Genetiker und Entomologe, in seinen Forschungsarbeiten mit den unterschiedlichen Farbgebungen des Birkenspanners (*Biston betularia*). In dieser Zeit waren bestimmte Regionen in England starken Industrieinflüssen ausgesetzt, wodurch in Folge erhöhter Emission von Fabrikgasen (Ruß und Schwefeloxid) eine Verdunkelung der Baumrinde resultierte. Kettlewell beobachtete ausgehend von einer Ursprungspopulation in diesen Gebieten, dass sich vor Allem die dunklere Form gegenüber der helleren Form des Birkenspanners behauptete.

Erklären Sie, wie die Dominanz der dunkleren Form des Birkenspanners in den Regionen mit industriellen Einflüssen evolutionsbiologisch zu Stande kam?

Die rußbedeckten Bäume boten keine ausreichende Tarnung mehr für die helleren Individuen der Birkenspanner. Aus diesem Grund wurden die hellen Individuen häufiger von ihren Fressfeinden erkannt und gefressen. Die Birkenspanner, welche durch eine Variation in der Ursprungspopulation stärker melanisiert waren, hatten einen Selektionsvorteil. Durch die dunklere Färbung, welche auf Mutation und Rekombination spezifischer Gene zurückzuführen ist, waren sie schlechter für Fressfeinde auf der veränderten Birkenrinde zu erkennen und wurden somit weniger stark dezimiert. Dieser Mechanismus führte dazu, dass die helleren Birkenspanner allmählich weniger wurden, wohingegen die stärker melanisierte Form einen größeren Anteil in der Population einnahm. In diesem Zusammenhang veränderte sich der Genpool der Birkenspannerpopulation in dieser Region, womit die Dominanz der dunkleren Individuen zu erklären wäre.

Beurteilung der Aufgabe 7

Bitte beurteilen Sie nun die Aufgabe 7 in der Klausur. Lesen Sie sich dazu die Antwort des Schülers/der Schülerin durch und tragen Sie anschließend in das Feld die Punktzahl ein, die Sie vergeben.

Beurteilung von Aufgabe 7 /20 Punkten

Frage zum Schülertext in Aufgabe 7:

Wie würden Sie die Antwort des Schülers in Aufgabe 7 beurteilen?

☐ Er ist wissenschaftlich korrekt.

☐ Er enthält eine anthropomorphe Fehlvorstellung.

☐ Er enthält eine teleologische Fehlvorstellung.

☐ Er enthält eine essentialistische Fehlvorstellung.

☐ Er enthält eine religiöse Fehlvorstellung.

Abbildung 4.5 Screenshot 3 of the SI

7. Anfang der 50er Jahre beschäftigte sich Bernard Kettlewell (1907-1978), ein britischer Genetiker und Entomologe, in seinen Forschungsarbeiten mit den unterschiedlichen Farbgenen des Birkenspanners (Biston betularia). In dieser Zeit waren bestimmte Regionen in England starken Industrieinflüssen ausgesetzt, wodurch in Folge erhöhter Emissionen von Fabrikabgasen (Ruß und Schwefeldioxid) eine Verdunkelung der Baumrinde resultierte. Kettlewell beobachtete aufgrund von einer Ursprungspopulation in diesen Gebieten, dass sich vor Allem die dunklere Form gegenüber der helleren Form des Birkenspanners behauptete.

Erklären Sie, wie die Dominanz der dunkleren Form des Birkenspanners in den Regionen mit industriellen Einflüssen evolutionsbiologisch zu Stande kam?

Die rußbedeckten Bäume boten keine ausreichende Tarnung mehr für die helleren Individuen der Birkenspanner. Aus diesem Grund wurden die hellen Individuen häufiger von ihren Fressfeinden erkannt und gefressen. Die Birkenspanner, welche durch eine Variation in der Ursprungspopulation stärker melaniert waren, hatten einen Selektionsvorteil. Durch die dunklere Färbung, welche auf Mutationen und Rekombination spezifischer Gene zurückzuführen ist, waren sie schlechter für Fressfeinde auf der veränderten Birkensrinde zu erkennen und wurden somit weniger stark dezimiert. Dieser Mechanismus führte dazu, dass die helleren Birkenspanner allmählich weniger wurden, wohingegen die stärker melanierte Form einen größeren Anteil in der Population einnahm. In diesem Zusammenhang veränderte sich der Genpool der Birkenspannerpopulation in dieser Region, womit die Dominanz der dunkleren Individuen zu erklären wäre.

Diagnostische Kompetenz bei Lehramtsstudierenden

Später fortfahren

Unfrage verlassen und Antworten löschen

Beurteilung der Aufgabe 7

Bitte beurteilen Sie nun die Aufgabe 7 in der Klausur. Lesen Sie sich dazu die Antwort des Schülers/der Schülerin durch und tragen Sie anschließend in das Feld die Punktzahl ein, die Sie vergeben.

Beurteilung von Aufgabe 7

19 /20 Punkten

Frage zum Schilertext in Aufgabe 7:

Wie würden Sie die Antwort des Schülers in Aufgabe 7 beurteilen?

☒ Er ist wissenschaftlich korrekt.

☐ Er enthält eine anthropomorphe Fehlvorstellung.

☐ Er enthält eine teleologische Fehlvorstellung.

☐ Er enthält eine essentialistische Fehlvorstellung.

☐ Er enthält eine religiöse Fehlvorstellung.

Abbildung 4.6 Screenshot 4 of the SI

5 STUDIE 2: THE SIMULATED CLASSROOM BIOLOGY (SCR^{Bio}) – A SIMULATED CLASSROOM ENVIRONMENT FOR CAPTURING THE ACTION-ORIENTED PROFESSIONAL KNOWLEDGE OF PRE-SERVICE TEACHERS ABOUT EVOLUTION⁴

Abstract:

The professional knowledge of pre-service teachers is highly important for effective and successful teaching. In recent years, many research groups have been engaged in developing simulated classroom environments to capture especially the pedagogical knowledge (PK) of pre-service teachers, neglecting the content-related facets of professional knowledge such as pedagogical content knowledge (PCK). Thus, validated test instruments are a prerequisite to obtain objective and reliable information about the action-orientated PCK of pre-service teachers and build on this to evaluate possible supporting measures. In the present study, we describe the development of a simulated classroom environment – the Simulated Classroom Biology (SCR^{Bio}) – and provide evidence regarding its validity to assess pre-service biology teachers' action-oriented PCK in the area of evolution. This study examined the evidence supporting the validity of using the SCR^{Bio} to investigate action-oriented PCK of pre-service biology teachers. The (1) evidence based on test content (expert ratings) and the (2) evidence based on relation to other variables (known-groups comparison) was obtained. We tested the SCR^{Bio} with $N = 76$ German pre-service biology teachers. Our results show the successfully operationalized PCK in the SCR^{Bio} through explicit allocation of specific misconceptions to each student answer and the valid measurement of pre-service biology teachers' action-oriented PCK compared to experimental groups.

Keywords:

Professional knowledge, Pre-service teacher, Simulated classroom environment, Evolution, Natural selection, Enacted pedagogical content knowledge (ePCK)

⁴ Fischer, J., Machts, N., Bruckermann, T., Möller, J., & Harms, U. (eingereicht 2021). Simulated Classroom Biology (SCR^{Bio}) - a simulated classroom environment for assessing the action-oriented professional knowledge of pre-service teachers about evolution. *Journal of Science Teacher Education*.

5.1 Introduction

Teachers strongly influence student achievement. Many scholars have shown that professional knowledge, the core of professional competence, is the key element for effective and successful teaching (e.g., Abell, 2007; Darling-Hammond, 2000; Hattie, 2009). Based on Shulman (1987), researchers generally agree that content knowledge (CK), pedagogical content knowledge (PCK) and pedagogical knowledge (PK) are essential domains of professional knowledge which influence teaching and learning (e.g., Abell, 2007; Baumert & Kunter, 2011). Within the conceptualization of professional knowledge in science education, Shulman (1986) stressed the relevance of PCK for creating learning opportunities and initiating student learning processes. Consequently, there has been a lot of research on what PCK is (e.g., Magnusson et al., 1999; Park & Oliver, 2008), how it differs from other knowledge domains (e.g., Großschedl et al., 2015; Großschedl et al., 2014; Mahler et al., 2017), and how it can be measured and supported (e.g., Henze & van Driel, 2015). Regarding PCK, knowledge about student understanding is essential for the diagnosis of students' misconceptions in the science teaching process and necessary for future students' learning success (Förtsch et al., 2018). Recently, a new conceptualization of PCK (Refined Consensus Model of PCK) was introduced by distinguishing different realms of PCK, of which the enacted PCK (ePCK) represents the specific knowledge and skills of teachers in particular teaching situations (Carlson & Daehler, 2019). Thus, the ePCK can be defined and described as a PCK close to action – the action-oriented PCK.

Previous studies on PCK lacked a focus on its application in teaching situations in the sense of capturing ePCK. The methodology of the most quantitative studies which record PCK through questionnaires are not suitable to measure ePCK (e.g., Brunner et al., 2006; Kleickmann et al., 2013; Großschedl et al., 2019). Qualitative approaches (for an overview see Depaepe et al., 2015), such as interviews (e.g., Jüttner & Neuhaus, 2013; Rozenszajn & Yarden, 2014), video and audio recordings (e.g., Hill et al., 2008; Park & Oliver, 2008) could address this problem and provide a deeper insight into ePCK. However, these methodologies often have small sample sizes and are very time-consuming.

Simulated classroom environments might be a solution for this problem, as they can capture PCK in action-oriented settings and can efficiently collect large samples at the same time. (e.g., Dalgarno et al., 2016). To date, most simulated classroom environments are developed to measure the teachers' PK in action (e.g., The Simulated Classroom: Fiedler et al., 2007; Fiedler et al. 2002; Kaiser et al., 2013; Kaiser et al., 2017; Südkamp et al., 2008), for example classroom management and diagnostic competence, but they neglect content-related domains (i.e., PCK, CK) of pre-service teachers' professional knowledge. Respective studies show that simulated classroom environments are suitable instruments for the assessment of action-oriented PK in teachers.

From these generic studies we conclude that simulated classroom environments represent appropriate tools to capture action-orientated, content-related professional knowledge (i.e., ePCK/CK) of pre-service teachers. To test this assumption we developed the Simulated Classroom Biology (SCR^{Bio}), a tool measuring (pre-service) teachers' content-related professional knowledge, i.e., ePCK and CK focusing on the central biological content evolution.

5.2 Theoretical Background

5.2.1 Teachers' Professional Knowledge

Professional knowledge is described as the core of professional behavior, a prerequisite for good teaching and predictive of student outcomes (e.g., Baumert & Kunter, 2006). Shulman's (1986) conceptualization of professional knowledge as CK, PCK, and PK has recently served as a framework for several large-scale studies that assessed and described professional knowledge in science (e.g., *ProwiN*: Fischer et al., 2012; *KiL*: Kleickmann et al., 2013) and in mathematics (e.g., *COACTIV*: Kunter et al., 2013). CK and PCK are mainly understood as the content-related domains of professional knowledge, whereas PK can be seen as a non-content-related (e.g., Förtsch et al., 2018; Jüttner et al., 2013; Mahler et al., 2017).

In general CK refers to knowledge of facts, terms and the conceptual understanding of these contents (Shulman, 1986) and is essential for effective teaching (e.g., Ball et al., 2001; Baumert et al., 2010; Friedrichsen et al., 2009) and student performance (Sadler et al., 2013). In general, the PCK domain comprises two facets. The first facet describes knowledge that the teacher uses to formulate and represent the subject matter in an intelligible way (Shulman, 1986) and is summarized as knowledge about instructional strategies. The second facet of PCK focuses on knowledge about students' understanding, which is mainly operationalized through the diagnosis of student misconceptions (Förtsch et al., 2018). In addition, individual research groups have added further facets, which are covered by the PCK, such as curricular knowledge and knowledge of assessment methods (e.g., Tamir, 1988). In contrast to the CK and PCK domain, the PK domain is not content-related and covers several facets, such as general knowledge (Kunter et al., 2013) and has a focus on general aspects of teaching (Shulman, 1987). Regarding PK, teachers' judgement accuracy is key and describes the relationship between teachers' judgements of students' achievement and actual student performance (Südkamp et al., 2012). For example, a student in biology class may answer some questions correctly or incorrectly. Here the teacher needs biological CK and PCK to make a diagnosis. The assessment of the student's performance and student's participation during the lesson requires teacher PK.

STUDIE 2

The judgement accuracy can be measured by three components, (1) the ranking component, which shows whether the teacher can correctly assess the characteristic expression of students, (2) the level component, which indicates the teachers' ability to correctly assess the average level of a student characteristic, and the (3) differentiation component, which shows the ratio of real performance to assessed performance (Südkamp et al., 2012).

A fundamental model of professional competence in the German-speaking countries is developed in the COACTIV project and includes cognitive (e.g., professional knowledge) as well as affectively motivational (e.g., self-regulation, motivational orientation) aspects (Kunter et al., 2011). However, this conceptualization neglects the action-oriented facets which Blömeke et al. (2015) integrated into a new model. The authors describe the observable teaching action of a person as shown performance. For a long time there was no differentiation between tacit and action-oriented knowledge within professional knowledge in science education, whereas following psychological approaches, types of knowledge include declarative knowledge (knowing, what) and action-related knowledge, consisting of procedural knowledge (knowing, how) and conditional knowledge (knowing, when and why; Paris et al., 1983). In 2017, the Refined Consensus Model of PCK (RCM) advanced among others, the framework of Shulman (1986) and the Magnusson Model (Magnusson et al., 1999). The RCM provides researchers with a differentiated analysis of teachers' PCK that goes beyond static knowledge and focuses on dynamic application in action-orientated situations. A core characteristic of this model is the differentiation of three distinct realms of PCK - the collective PCK (cPCK), the personal PCK (pPCK), and the enacted PCK (ePCK) – ranging from the professional knowledge of the community of educators in a field (cPCK), the personal professional knowledge of individual science teachers (pPCK), to specific knowledge and skills of teachers in particular teaching situations (ePCK). Still, the model accounts for the domains of teachers' professional knowledge (i.e., content knowledge, pedagogical knowledge, knowledge of students, curricular knowledge and assessment knowledge) from previous models (e.g., Magnusson et al., 1999; Shulman, 1987). The categorization of the PCK into three different knowledge realms (cPCK, pPCK and ePCK) provides a clear theoretical framework for the research within PCK (Carlson & Daehler, 2019). The RCM facilitates the assessment of different PCK domains - including the knowledge of students' understanding - in an action-oriented framework (ePCK).

Enabling pre-service teachers to apply their static knowledge in specific teaching situations and measuring ePCK simulated classroom environments might be a successful approach (Hixon & So, 2009).

5.2.2 Simulated Classroom Environments

Simulated learning environments have recently become increasingly popular in the context of teacher education (Bradley & Kendall, 2014). Simulated classrooms provide solutions to bridge the gap between real teaching situations in the classroom at school and theory bound teacher education at university (Hixon & So, 2009). A simulation is a simplified but valid, accurate and dynamic model of reality implemented in a digital system (Sauvé et al., 2007). New digital technologies have motivated many research groups to design classroom simulations and provide them as a tool for observing pre-service teachers' application of their professional knowledge under controlled conditions (e.g., Cheong, 2010; Gurvitch & Metzler, 2009). Most classroom simulations in teacher education focus on the facets of PK, including classroom management, self-efficacy and the professional interpersonal competence (e.g., *TeachMETM*: Bautista & Boon, 2015; *simSchool*: Christensen et al., 2011; *VirtualPREX*: Dalgarno et al., 2016; *CS-TGCTS*: Yeh, 2004) and neglect the content-related facets like PCK. Here, the computer-based simulated classrooms facilitate remote access role-plays of the pre-service teacher and virtual students, in which the teacher faces real, classroom-specific problems (Clapper, 2010). In a role-play simulation the user slips into the role of the individuals involved in teaching (teacher or student) and adopts their tasks in this simulated scenario (Veletsianos et al., 2010). Generally, role-play simulations can be distinguished into the non-immersive simulation which uses texts and images to create a classroom situation and the immersive simulation which provides visual representations of a classroom including the application of avatars (Dalgarno et al., 2016). The realistic integration of the participants through 3-D representations in the classroom by a role-playing avatar enables them to identify more closely with their assigned teacher role (Jamaludin et al., 2009). In contrast, non-immersive classroom simulations have the advantage of providing complexity reduced environments under controllable conditions (e.g., Fiedler et al., 2007; Südkamp et al., 2008).

Non-immersive simulation, like the Simulated Classroom (SCR) is a suitable tool to illustrate classroom interaction (i.e., role play between teacher and students) in an action-oriented setting. Here, the pre-service teachers slip into the role of the teacher and can pose questions to the virtual students in a virtual lesson. Additionally, the SCR can capture different facets of professional knowledge, which must be applied by the teacher to achieve an accurate assessment of the virtual students. Initially, the SCR was a computer-based tool to measure action-oriented diagnostic knowledge for the PK domain in an experimental environment on judgement biases such as teachers' error of central tendency and naïve theories about the behavior of individuals and stereotypes (Fiedler et al., 2007; Fiedler et al., 2002). Südkamp et al. (2008) expanded the existing SCR in order to handle more complex educational psychological contents. Now the researchers could integrate their own tasks, questions and answers into the system, which in turn increased the ecological validity

of the instrument. This made it possible to analyze the diagnostic judgements of pre-service teachers in a more differentiated way. Recently, the SCR was used to disentangle effects of virtual students' achievement and motivation on teachers' assessments of both constructs (Kaiser et al., 2013) and to analyze how minority status impacts teachers' assessments (Kaiser et al., 2017).

Whilst in educational psychology teachers' accuracy of judgements, based on the differentiation, ranking, and level-component (Fiedler et al., 2002; Fiedler et al., 2007; Südkamp et al., 2008) is a measure of diagnostic knowledge (i.e., PK), models on teachers' PCK consider the diagnosis of subject-specific misconceptions as key (e.g., Großschedl et al., 2018; Ziadie & Andrews, 2018). Teachers need to know valid and reliable assessments to identify students' misconceptions in their subject (e.g., CANS, CINS, ACORNS for evolution in biology; Ziadie & Andrews, 2018). However, teacher-student interactions require consistent diagnosis of students' misconceptions (i.e., ePCK) during real classroom discussions on e.g., natural selection. In order to assess pre-service teachers' knowledge about students understanding during student-teacher-interactions, the initial SCR was expanded to the Simulated Classroom Biology (SCR^{Bio}), which focuses the theory of evolution.

In previous studies, the SCR was used to measure action-oriented PK (Fiedler et al., 2002; Fiedler et al., 2007; Südkamp et al., 2008). However, there is a lack of evidence on what findings need to be considered in research on teaching and learning evolution within biology education by developing a simulated classroom environment and, whether the ePCK of pre-service teachers in a simulated classroom (i.e. through the SCR^{Bio}) leads to valid conclusions.

5.2.3 Teaching and Learning Evolution

Within biology education, the diagnosis of students' understanding about the evolutionary theory is paramount because evolutionary theory is key to a conceptual understanding of all topics in the field of biology (e.g., Anderson et al., 2002; Basel et al., 2014; Furtak, 2012; Opfer et al., 2012). Unfortunately, the evolutionary theory has always been considered very challenging in the science curriculum, and traditional teaching often fails when it comes to expanding students' understanding to scientific explanations (e.g., Bishop & Anderson, 1990; Gregory, 2009; Kampourakis & Zogza, 2008; McVaugh et al., 2011; Opfer et al., 2012; Shtulman, 2006). Educators and researchers identified common misconceptions among students as well as the public and instructional strategies to overcome these. Thus, a knowledge base is provided for the professional knowledge of teachers about the theory of evolution (for an overview: on misconceptions, see Gregory, 2009; on teaching strategies, Ziadie & Andrews, 2018; Harms & Reiss, 2019).

Despite its functionality as conceptual organizer for learning about evolution (e.g., Opfer et al., 2012; Tibell & Harms, 2017) the key concept of natural selection itself evokes numerous misconceptions, especially in explaining the adaption of individuals and population to their environment (e.g., Bishop & Anderson, 1990; Ferrari & Chi, 1998). Most misconceptions are based on the belief that evolution occurs because individual organisms actively evolve and adapt during their lifetime (e.g., Bishop & Anderson, 1990; Settlage, 1994; Nehm & Schonfeld, 2008). Teleological and anthropomorphic explanations as well as the misconception of use or disuse are widespread and intensively researched categories of misunderstanding. In teleological explanations the development of a trait is target-oriented and purposeful. Here, traits arise because they result in a survival advantage for the organism (e.g., Alters & Nelson, 2002; Andrews et al., 2011; Bishop & Anderson, 1990; Kampourakis & Zogza, 2008; Nehm et al., 2009; Nehm & Reilly, 2007; Nehm & Schonfeld, 2008; Settlage, 1994; Sinatra et al., 2008). Anthropomorphic misconceptions are characterized by the transfer of human characteristics – physical and mental – on other living organisms (Byrne et al., 2009). Trait changes result from intentional and purposeful action through the individual to cope with new environmental conditions (e.g., Demastes et al., 1995; Gregory, 2009; Sinatra et al., 2008; Tamir & Zohar, 1991). Similar to anthropomorphic misconceptions is the belief that individuals also have an influence on the physical change through their use or disuse of organs or abilities. After the development of the new traits these are passed on to future generations (e.g., Andrews et al., 2011; Kalinowski et al., 2016; Kampourakis & Zogza, 2008; Nehm et al., 2009; Prinou et al., 2008). All these misconceptions ignore several evolutionary mechanisms of natural selection and follow a causal explanatory pattern. Supporting students in acquiring conceptual knowledge about the theory of evolution requires teachers to diagnose the most common student misconceptions (e.g., Gregory, 2009), which the teachers' PCK on the theory of evolution presupposes (Ziadie & Andrews, 2018).

Different contexts influence the probability of an expressed misconception regardless of the specific misconception categories. For example, the biological domain (i.e., bacteria, plants, animals, humans) influences the appearance of a misconception (Opfer et al., 2012). For example, there are fewer misconceptions in connection with plants than with animals (Nehm et al., 2012). Moreover, students are more adept at explaining a gained trait than the loss of a trait in the process of natural selection (Nehm et al., 2012; Nehm & Ha, 2011).

5.3 Research Questions

In this study we investigate the validity evidence for the use of the SCR^{Bio} to measure pre-service biology teachers' knowledge about student understanding (here: diagnosis of

STUDIE 2

specific virtual student misconceptions regarding evolution) in an action-oriented setting (ePCK). We consider (1) evidence based on the test content and (2) evidence based on known-groups comparisons as sources of validity evidence.

The SCR^{Bio} should provide a measure of pre-service teachers' ePCK, which is reflected in the diagnosis of misconceptions. The virtual students' answers must be reliably assigned to the operationalized misconceptions. Therefore, we used expert ratings to verify the content of the virtual student answers, i.e., whether the answer actually contains a specific misconception or expresses a scientific way of thinking.

- I. To what extent do the expert judgements support the validity evidence of the developed virtual student answers with regard to the specific categories of misconceptions (*validity evidence based on test content*)?

For the measurement of pre-service teachers' ePCK, scores on the measure should vary for groups known to differ in terms of the specific information prompts. We therefore created groups with specific PCK knowledge by prompting PCK-related or PCK-unrelated information.

- II. To what extent is the SCR^{Bio} sensitive to the given PCK-related prompt in biology education (PCK) and educational psychology (PK) (*validity evidence based on known-groups comparison*)?

Hypotheses

- I. Experts who have not been involved in the item development can correctly judge the specific category of misconception or scientific explanations that the items in the SCR^{Bio} are intended to measure.
- II. The SCR^{Bio} has good discriminatory power, that is, a PCK-related prompt:
 - a. increases pre-service teachers' diagnoses of specific misconceptions categories in virtual student answers (refer to PCK).
 - b. does not increase pre-service teachers' judgment accuracy of virtual students' lesson achievement (refer to PK).

5.4 Methods

5.4.1 Sample

The participants were $N = 76$ pre-service teachers from two universities in Northern Germany enrolled in a teacher education program (24% male) with biology as one of their subjects. The teacher preparation in Germany is organized in two phases with phases leading to a degree after three years (bachelor) and five years (master). A total of $n = 38$

participants studied in the Bachelor teacher education program, $n = 37$ in the Master teacher education program and one participant did not provide any information. A proportion of 86.8% aspired to a teaching profession for the academic track (i.e., grammar school) and 13.2% for the nonacademic tracks (i.e., interdenominational school). The age of the study participants ranged between ages 18 and 42, mainly between ages 21–23 (51%), 24–26 (26%), and 27–30 (16%).

5.4.2 Development of the Simulated Classroom Biology (SCR^{Bio})

The initial developed non-immersive Simulated Classroom (Fiedler et al., 2002; Fiedler et al., 2007; Südkamp et al., 2008) was expanded to the SCR^{Bio}. Therefore, we created a virtual biology lesson for the application in the SCR^{Bio} with questions on evolution. We prepared ($N = 27$) questions on the process of natural selection covering all groups of organisms (i.e., bacteria, plant, animal, and human). Furthermore, we developed the questions to address the gain or loss of a trait as well as whether the trait has a morphological or physiological origin. For each individual question we designed possible virtual students answers ($N = 476$; between 8–30 words), based on theory and authentic student answers from exams (Baalmann et al., 2004; Gregory, 2009; see Table 5.1). For each question ($N = 27$; questions having between 30–84 words) the virtual students' answers represented a scientifically correct explanation or a misconception (anthropomorphic, teleological, use or disuse; see Table 5.2).

Tabelle 5.1 Scheme for the development of evolutionary biological questions and possible student answers in the SCR^{Bio}.

Evolutionary Direction	Type of Trait	Domain			
		Bacteria	Plant	Animal	Human
Loss	Morphological	2 questions with 35 answers	3 questions with 50 answers	3 questions with 54 answers	2 questions with 28 answers
	Physiological				
Gain	Morphological	5 questions with 93 answers	4 questions with 74 answers	4 questions with 76 answers	4 questions with 66 answers
	Physiological				

STUDIE 2

Tabelle 5.2 Example question and possible student answers about the red bentgrass (botanical organism, gain of a trait, physiological trait)

Red Bentgrass (<i>agrostis tenuis</i>)
Possible teachers' question: <i>"The red bentgrass is a common grass sometimes found close to copper-mines. In these locations the soil is very coppery and has a toxic effect on many plants. Investigations show that the red bentgrass shows a high copper-tolerance that could not be established in conspecifics living beyond the coppery soils. How can the tolerance to copper be explained considering principles of evolutionary biology?"</i>
Anthropomorphic students' answer: <i>"The red bentgrass had the intention to develop a copper-resistance in order to be able to live on soils around copper-mines."</i>
Teleologic students' answer: <i>"The copper-resistance of the grass had the purpose to grow even in areas around copper-mines."</i>
Use or disuse students' answer: <i>"The intense utilisation of the copper-resistance on polluted soils accelerated the development of this trait."</i>
Scientific students' answer: <i>"The change of the environmental conditions led to the shift of the evolutionary pressure whereby only those plants that are copper-resistant can survive in coppery soils."</i>

5.4.3 Procedure

For this study, the pre-service teachers taught a virtual lesson within the computer-based simulation SCR^{Bio} on laptops. The laptops were set up in university lecture halls and pre-service teachers' participation in the study lasted 60 minutes. In the virtual lesson, the pre-service teachers performed a teaching sequence—following a question-answer scheme to diagnose virtual students' misconceptions—and a grading sequence—to assess the overall performance of the class. The simulation of the virtual lesson adhered to the procedure described below.

Before starting the teaching and grading sequence, the SCR^{Bio} on the laptop screen informed the pre-service teacher of the duration of the study, information about data privacy, the virtual lesson's topic (evolution - natural selection) and the class level to be taught (12th grade). When the SCR^{Bio} simulation started, the pre-service teachers were randomly assigned to either a control- or the experimental group. In the control group receiving PCK-unrelated information prompts, pre-service teachers were given a brief overview on biology in school, science and education. In the experimental group receiving PCK-related information prompt, pre-service teachers were given an overview of the three misconception categories and a scientific explanation.

The teaching sequence of the virtual lesson started after the pre-service teachers received the different information prompts. The teaching sequence followed a question-answer scheme in which the pre-service teacher asks a question and the virtual students

answer. The pre-service teachers could ask these virtual students questions about natural selection that could be selected from a set of pre-defined questions (see Development section). In the SCR^{Bio} the virtual class consisting of nine virtual students represented by students' photos and their names. The virtual students' answers varied according to their pre-set ability profile. The class was divided into three groups of three virtual students. For each group a different ability profile was determined (i.e., 20%/50%/80% probability of correct answers). If the virtual students answered incorrectly, they always expressed a specific misconception about evolution (i.e., anthropomorphic, teleological or use/disuse). We combined each specific misconception category with each ability profile, what resulted in the nine students. For example, one of the virtual students answered only 20% correctly, while all other answers articulated a teleological misconception. Students who wanted to answer a question virtually raised their hand. This was indicated by a little hand-icon and orange-colored background of the students' photo (see virtual students 1–7 in Figure 5.1). Pre-service teachers then selected one virtual student to answer the question. The virtual student's answer was provided in a box. The pre-service teacher immediately had to diagnose the answer after reading it. This means that the pre-service teachers had to assess whether the virtual students' answer was scientifically correct or not and, if it was wrong, which misconception was expressed. The pre-service teachers had 30 minutes to ask as many questions and call up as many virtual students as they needed to diagnose each virtual student's misconception and overall class performance regarding natural selection. The grading sequence immediately followed.

In the grading sequence, the pre-service teachers had to assess the virtual students' proportion of correct answers (on a scale from 0–100%), which is described as the overall performance (i.e., lesson achievement). Finally, pre-service teachers had to enter their demographical data and were informed about the aims of the SCR^{Bio}.

Simulated Classroom Biology Token: xEcvYW Time: 11:36

Virtual class: Presentation of the nine students and their names. The students who want to give an answer are highlighted in orange and marked with a little finger symbol.

Category	Subcategory	Questions	Question addressed to class
Questions	Animal	Oenothera biennis - toxins	The red bentgrass is a common grass sometimes found close to copper-mines. In these locations the soil is very coppery and has a toxic effect on many plants. Investigations show that the red bentgrass shows a high copper-tolerance that could not be established in conspecifics living beyond the coppery soils. How can the tolerance to copper be explained considering principles of evolutionary biology?
	Plant	Cuscuta - plant organs	
	Human	Tobacco plant - defense mechanisms	
	Bacteria	Stomata in land plants	
		Mimulus - flower morphology	
		Red bentgrass - copper resistance	Question addressed to the class - complete question: The complete question which the students should answer is displayed.
		Trifolium repens - cyanide production	
		Questions - Plant species: Seven questions can be selected for the group of plants (here: the copper-resistance of the red bentgrass).	

Category - Set of questions: The following questions can be selected by the pre-service teachers.

Subcategory - Groups of organisms: The pre-service teachers could ask questions about the natural selection of bacteria, plants, animals, and humans (here: question on natural selection of plants).

Abbildung 5.1 Screenshot of the teaching sequence in the SCR^{Bio}.

5.4.4 Measures

During the teaching sequence we captured the pre-service teachers' assessment of the virtual students' misconceptions and scientific explanations by recording the number of accurate diagnoses as a measure of their ePCK. During the grading sequence we recorded the pre-service teachers' evaluation of the proportion of virtual students' correct answers over the entire virtual lesson as a measure of their PK. All data were recorded and stored as csv-data sets.

ePCK: We operationalize ePCK as the knowledge about student understanding. This is reflected in the diagnosis of specific categories of misconceptions in virtual students' answers by the pre-service teachers during the teaching sequence in the SCR^{Bio}. The relative frequency of correct diagnosis of misconceptions in virtual student answers is quantified as a measure of ePCK.

PK: We operationalize PK as the judgement accuracy regarding the assessment of the virtual students' lesson achievement by the pre-service teachers during the grading sequence in the SCR^{Bio}. We measure judgement accuracy in three components (Südkamp et al., 2012):

- i. The ranking component, as the correlation coefficient between pre-service teachers' ratings of virtual students' relative frequency of correct answers and their actual relative frequency.
- ii. The level component, as the difference between the class mean of pre-service teachers' ratings of virtual students' relative frequency of correct answers and the actual class mean

- iii. The differentiation component, as the ratio of the class variance of pre-service teachers' ratings of virtual students' relative frequency of correct answers and the actual class variance

5.4.4 Validity

In the development and evaluation of a test instrument to measure teachers' PCK validity is a fundamental element to be considered (e.g., Großschedl et al., 2019). Validity indicates the degree to which theory and evidence support the interpretation of test results (Kane; 2013). Accordingly, validation concerns not the test instrument itself but the interpretation of results that the test provides (AERA et al., 2014).

Validity Based on the Test Content

Expert judgements give an indication of the content validity between the test items and the theoretical construct (AERA et al., 2014; Sireci & Faulkner-Bond, 2014). The experts ($N = 10$) were mainly from the field of biology education and included both researchers and practitioners with various degrees in biology. They evaluated the items in terms of misconceptions by assessing each item's belonging to one of the three categories (i.e., anthropomorphic, teleological, use or disuse) or scientifically correct. Within SCR^{Bio} , ePCK is measured by diagnosing the misconceptions articulated in the answers of the students. Accordingly, it is mandatory that our student answers explicitly express the intended misconception. Each expert received a coding manual which in theory clearly defines the misconceptions to be judged and presents a number of examples. The content validity ratio ($CVR = \left(n_I - \left(\frac{N}{2} \right) \right) / \frac{N}{2}$) by Lawshe (1975) was adapted to quantify the agreement between the expert judgements. Here n_I = number of experts who judge an item as theoretically intended (i.e., the misconception we integrated into the virtual student answer was successfully identified by the experts) and N = total number of experts ($N = 10$). A CVR of at least 0.78 is necessary to consider an item or scale as valid (Frey, 2018).

Due to the large number of items (i.e., questions and corresponding answers), each expert rated a baseline pool ($N = 8$) and an additional share of the remaining questions ($N = 19$) and answers. The experts rated eight questions (baseline pool; questions from the different groups of organisms: bacteria, plants, animals and humans) and the corresponding 128 possible virtual student answers. In the second part, the experts were divided into five groups (two experts per group) which had to evaluate another four questions and corresponding virtual student answers ($N = 64$) in addition to the baseline pool. We were able to cover all 27 questions and corresponding student answers for the SCR^{Bio} with this expert rating scheme. We revised or selected the questions and the corresponding virtual student answers based on the feedback from the experts.

Validity Based on Relation to Other Variables (Known-group Comparison)

The known-groups method requires that at least two different groups work on a test. Thus, the test values can be interpreted as valid if the two groups show different, theoretically predicted results with respect to the same construct, depending on their preconditions (Hattie & Cooksey, 1984). This approach counts as a validity criterion to interpret the results of the construct – knowledge about student understanding – in the SCR^{Bio} validly. The sample is divided into two groups with different information prompts to provide evidence based on a known-groups comparison. The relative diagnosis rates for the identification of specific misconceptions and of scientific explanations in virtual student answers were calculated separately for both groups. A Pearson Chi-square test was performed to show group differences of these rates. Differences indicate the extent to which the information influences the diagnosis of misconceptions and scientific explanations within biology education (ePCK). Additionally, we addressed the components of judgement accuracy between the groups. We applied t-tests to determine if the PCK prompt influenced judgement accuracy of the pre-service teachers as a measure of pedagogical knowledge (PK).

5.5 Results**RQ1: Validity Evidence Based on Test Content**

The experts evaluated the baseline pool of virtual student answers with an overall $CVR_{mean} = 0.915$. The CVR_{mean} exceeds the threshold of 0.78 (Frey, 2018), indicating that virtual student answers clearly expressed a specific misconception or scientific explanation. A differentiated view of the expert rating shows the items (baseline pool) were evaluated with an overall $CVR_{mean} = 0.911$ and the remaining sets of items were evaluated with an overall CVR_{mean} between 0.859 and 1. Furthermore, the items were assessed regarding evolutionary direction of the trait (i.e., gain or loss) with an inter-rater agreement of 96.25% and the type of a trait (i.e., morphological or physiological) with an inter-rater agreement of 86.25%. This also confirmed the conceptual development of our question pool.

RQ2: Validity Based on Relation to Other Variables (Known-group Comparison)

Pre-service teachers in the control group who received PCK-unrelated information asked a total of $n = 1378$ questions, whereby 50.7% ($n = 699$) of the virtual student answers contained a misconception. They correctly diagnosed anthropomorphic misconceptions in 43.0% ($n = 99$), use or disuse misconceptions in 56.7% ($N = 123$), and teleological misconceptions with 29.4% ($n = 74$) (see Figure 2). Overall, the control group selected the appropriate category across misconceptions with 43.0% ($n = 300$; for a graphic representation see Figure 5.2).

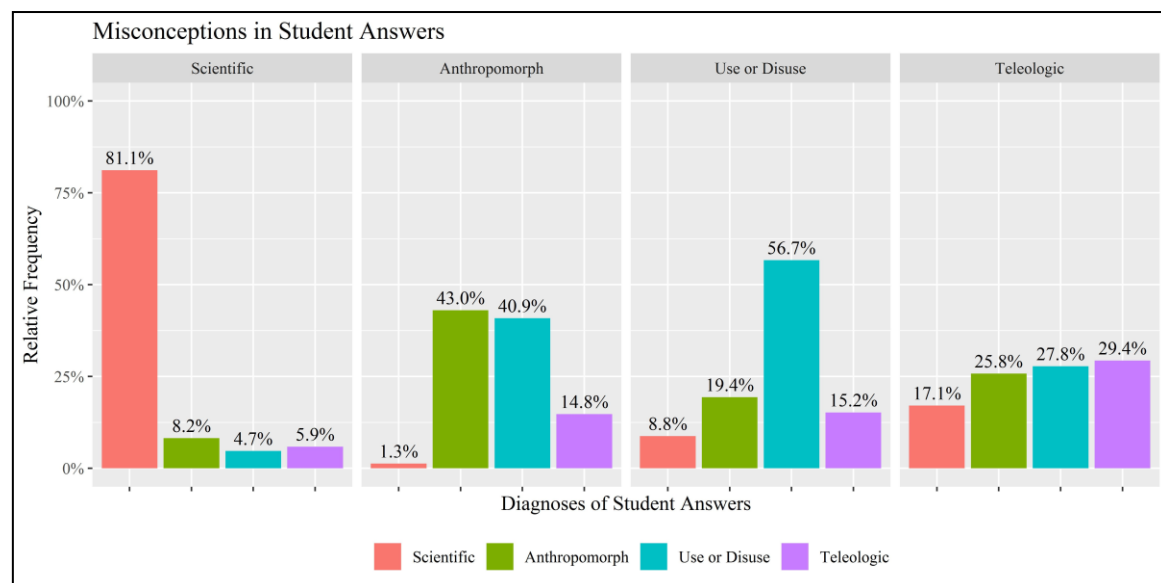


Abbildung 5.2 Control group: Relative frequency of diagnosis of students' answers with regard to expressed misconceptions categories or scientific reasoning.

In contrast, pre-service teachers in the experimental group who received PCK-related information asked a total of $n = 1556$ questions, whereas 51.2% ($n = 796$) virtual student answers were scientifically correct, and 48.8% ($N = 760$) expressed a misconception. They correctly diagnosed 76.2% ($n = 198$) of anthropomorphic misconceptions, 70.8% ($n = 177$) of misconceptions regarding use or disuse, and 56.4% ($n = 141$) of teleological misconceptions. In total, they correctly diagnosed 67.8% of all misconceptions and showed a significantly higher rate than the pre-service teachers in the control group ($\chi^2(1) = 97.28, p < .001$). Regarding the correct diagnosis rate of the scientific answers of the virtual students, the difference between both groups was not significant ($\chi^2(1) = 0.06, p = .804$) (for a graphic representation see Figure 5.3).

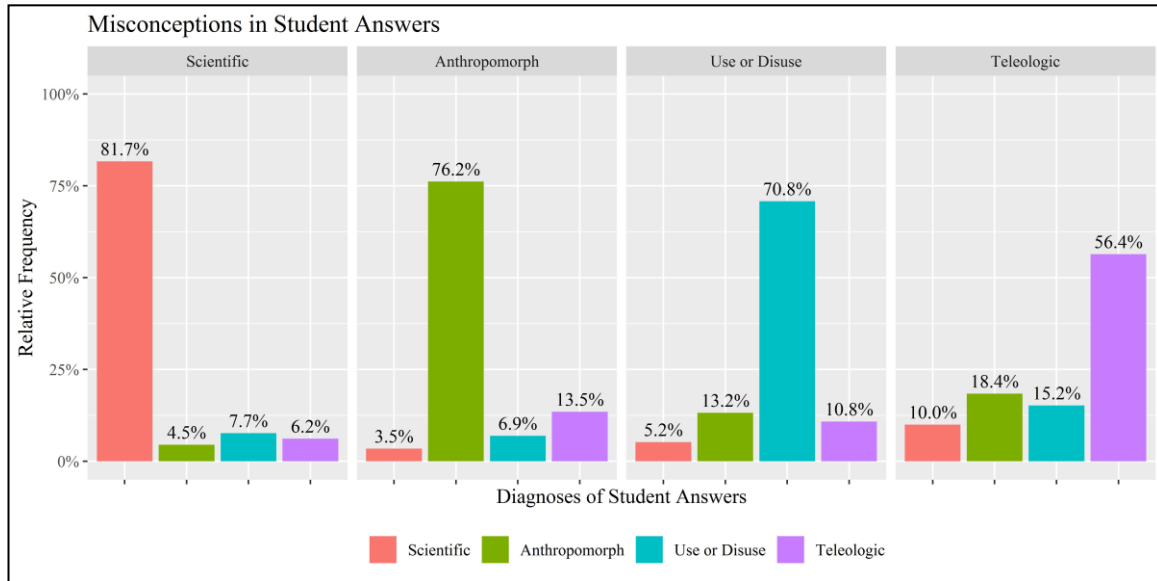


Abbildung 5.3 Experimental group: Relative frequency of diagnosis of students' answers with regard to expressed misconceptions categories or scientific reasoning.

We examined the differences between the experimental and the control group on the three judgement accuracy components to check whether the PCK-related information has an impact on the judgement accuracy of the pre-service teachers. Pre-service teachers in the control group reached a mean ranking component of $M_r = .31$ ($SD = .34$, $Min = -.48$, $Max = .88$), whereas the experimental group has a mean ranking component of $M_r = .34$ ($SD = .35$, $Min = .30$, $Max = .90$), the difference was not significant, even with an increased alpha-level of a $\alpha = .25$ ($t(73) = 0.27$, $p = .787$). For the mean level component, a result of $M_n = -.02$ ($SD = .17$, $Min = -.51$, $Max = .38$) was obtained for the control group and $M_n = .05$ ($SD = .16$, $Min = -.35$, $Max = .33$) for the experimental group. However, the corresponding t-test does not allow rejecting group difference in the level component of judgment accuracy ($t(73) = 1.94$, $p = .056$). Within the control group there is a mean differentiation component of $M_d = .57$ ($SD = .25$, $Min = .04$, $Max = 1.10$) and for the experimental group of $M_d = .60$ ($SD = .33$, $Min = .06$, $Max = 1.61$). The t-test shows no significant group differences with respect to the differentiation component ($t(73) = 0.47$, $p = 0.643$; see Table 5.3).

Tabelle 5.3 Results of the ranking, level, and differentiation component subdivided by control and experimental group.

	Control group				Experimental group			
	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>
Ranking component	.31	.34	-.48	.88	.34	.35	.30	.92
Level component	-.02	.17	-.51	.38	.05	.16	-.35	.33
Differentiaion component	.57	.25	.04	1.10	.60	.33	.06	1.61

Sample. $N = 76$.

5.6 Discussion

The present study reports the validity of using a German-language simulated classroom environment called the SCR^{Bio} to investigate and measure the ePCK of pre-service biology teachers. More precisely, it reports the diagnosis of common misconceptions about the evolutionary process of natural selection. Thus, the SCR^{Bio} must be developed in a way that the construct (i.e., knowledge about student understanding) is detected without the influence of construct-irrelevant variance (AERA et al., 2014). We focus on validity evidence based on test content and relationships to other variables.

The first research question is addressed using the validity evidence based on the test content gained by expert judgements. The test content of the items needs to adequately represent the construct knowledge about student understanding which is necessary for the valid interpretation of the test results. In this study researchers working in the field of biology education and biology teaching assigned the items (baseline pool) to common misconception categories or scientific ways of thinking. This resulted in an almost perfect agreement ($CVR_{\text{mean}} = 0.911$; see Ayre & Scally, 2014) with the intended wording of the items. Additional items have similar agreements (CVR_{mean} between 0.859 - 1). However, the overall CVR_{mean} of 0.915 is proposed to be sufficient for indicating validity based on test content (Ayre & Scally, 2014). Thus, the virtual student answers clearly expresses a specific category of misconception or a scientific way of thinking and we can assume the virtual student answers can be used as a valid measure of the knowledge about student understanding. Confirmation biases can be ignored (Guion, 1977) as the experts were not involved in the development of the SCR^{Bio}.

The second hypothesis, which focuses on the validity criterion of known groups, predicted that pre-service teachers who received a PCK-related information prompt (experimental group) show a higher diagnosis of specific misconception categories than their colleagues who received a PCK-unrelated information prompt (control group). In

accordance with the hypothesized group differences, we found a significant difference in the diagnostic rate of the specific misconception categories between the groups ($\chi^2(1) = 97.28, p < .001$). Thus, the PCK-related information prompt influences the diagnosis of misconceptions positively and revealed the sensitivity of the SCR^{Bio} for ePCK. In this regard, the diagnosis of misconception is based on the ePCK of the pre-service teachers (e.g., Förtsch et al., 2018; Lucero et al., 2017). We conclude that this evidence supports the validity of our interpretation of the test results as measures for ePCK since the group means differ in the expected directions. In contrast, we observed no group differences in the diagnosis of scientifically correct virtual student answers ($\chi^2(1) = 0.06, p = .804$), indicating that the knowledge for the content-related assessment of an answer (i.e., scientifically correct or incorrect; refers to CK) differs from the knowledge about student understanding (i.e., diagnosis of specific misconception categories, refers to PCK; e. g., Förtsch et al., 2018; Großschedl et al., 2015). The PCK-related information prompt does not seem to influence the accuracy of pre-service teachers (based on their PK, Südkamp et al., 2008), in terms of the correlation between judgments of lesson achievement and actual achievement. Neither does it seem to influence the lesson judgment variance relative to the actual achievement variance. However, we cannot exclude that the PCK-related information on misconceptions led to different mean class judgment levels, possibly in terms of stricter judgments in the experimental condition. Thus, these results provide clues to the distinction of knowledge needed to (1) diagnose common misconceptions in virtual student answers (ePCK) and (2) assess virtual students' lesson achievement (PK).

As previously mentioned, validity cannot be stated universally but rather only for a specific interpretation of test results (AERA et al., 2014; Kane, 2013). In summary, this study provides evidence that the content of the items corresponds to the theoretical meaning of knowledge about student understanding (evidence based on test content) and indicates the sensitivity of the SCR^{Bio} for ePCK by differences in the relative frequency of diagnosing specific misconceptions categories in the virtual students answers between the experimental and control group (known groups). Therefore, the present simulated classroom environment – the SCR^{Bio} – may be validly used to provide findings of the ePCK in biology education. Furthermore, the validation of results in the SCR^{Bio} contributes to closing the gap from the availability of simulated classroom environments, which only cover the pre-service teachers' action-oriented PK by now allowing the measurement of ePCK in biology.

5.6.1 Limitations

The newly developed SCR^{Bio} is the first attempt to investigate ePCK of pre-service teachers in an action-oriented simulated classroom environment. In accordance with our study design we used the following sources for validity evidence: (1) evidence based on the test

content and (2) evidence based on relation to other variables to provide a use argument for the valid interpretation of the test results in the SCR^{Bio}. However, the *Standards for Educational and Psychological Testing* provide various further sources of validity. These include the evidence based on internal structure, evidence based on response processes, and evidence based on the consequences of testing (AERA et al., 2014), which we did not include in this first validation approach.

Our approach for providing content validity was relatively straightforward compared to other approaches. We used a distinct operationalization of misconceptions which likely led to the very high agreement on the answer categories across expert raters. In the SCR^{Bio} we used virtual student answers that are somewhat simplistic compared to more potential real-life student answers.

The known-groups comparison in the SCR^{Bio} was based on an information prompt in which a part of the sample received information helpful in processing the SCR^{Bio} while the remaining pre-service teachers received no helpful information. This comparison supports the valid interpretation of the results in the SCR^{Bio}. However, other group differences, which we have not considered in order to support this validity criterion, can also be taken into account, such as differences regarding the teaching career (e.g., academic or non-academic track); the status of the university career (e.g., graduate or undergraduate students), the teaching experience (e.g., pre-service or in-service teachers) or the studied subjects (e.g., one or two natural science subjects; e.g., Großschedl et al., 2019; Hartmann et al., 2015; Mathesius et al., 2019).

The SCR^{Bio} is a complexity reducing simulated classroom environment designed to come close to reality of imitate a real classroom (e.g. Fiedler et al., 2007; Südkamp et al., 2008). However, the instrument is still an environment of reduced information, blanking out other factors that influence teaching, such as student disruptions or answers that do not represent attempts to successfully answer the posed question (i.e., classroom management; PK). Thus, it is necessary to control the manipulated variables in the SCR^{Bio} to reduce the noise of the investigation.

Finally, in the current SCR^{Bio}, we used student answers from three common misconception categories (i.e., anthropomorphic, teleological and use or disuse). However, research has identified various categories of misconceptions about evolution (e.g., Bishop & Anderson, 1990; Gregory, 2009; Settlage, 1994; Nehm & Schonfeld, 2008). For example, religious beliefs were abandoned because they do not have much influence in Germany (e.g., Großschedl et al., 2014), but in other countries they do (e.g., Berti et al., 2010; Billingsley et al., 2016; Rissler et al., 2014).

5.6.2 Implications for Further Research

Few simulated classroom environments have been developed to measure and promote action-oriented facets of PK (e.g., Bautista & Boon, 2015; Christensen et al., 2011; Dalgarno et al., 2016), but none that cover ePCK. We were successful in validating the SCR^{Bio} in German so far, but still aim to use and compare the instrument in other countries and expand the range of participants. We therefore integrated new questions about the process of natural selection and corresponding answers from American college students (Nehm et al., 2012; *ACORNS*). Future studies with the SCR^{Bio} will focus on the extent to which pre-service teachers are able to diagnose misconceptions with regard to the evolutionary direction of a trait (i.e., loss or gain), the type of trait (i.e., physiological or morphological) or the domain (i.e., bacteria, plant, animal, human) of the biological example. Some studies have already examined the influence of these factors in relation to diagnostic ability (e.g., Nehm et al., 2012; Opfer et al., 2012). Beyond the diagnosis of misconceptions (e.g., Gregory et al., 2009), pre-service teachers need to cover another key facet of PCK - instructional strategies (e.g., the conceptual change strategy) - in order to respond adequately to misconceptions regarding evolution (e.g., Harms und Reiss, 2019; Ziadie & Andrews, 2018). New versions of the SCR^{Bio} will integrate adequate instructional strategies (i.e., strategies to overcome specific misconceptions), which must be selected by the pre-service teacher. We are currently conducting initial validation studies. In the future the SCR^{Bio} will be developed from an assessment instrument to a training tool to simulate explicit teaching situations. In this training tool university students will have the opportunity to train action-oriented professional knowledge. A further goal is to provide direct feedback on performance in the SCR^{Bio} through a new developed report module. Thus, the SCR^{Bio} will significantly contribute to preparing future biology teachers for the challenges of teaching and learning evolution in class.

5.7 References

- Abell, S. K. (2007). Research on science teacher knowledge. In S. K. Abell (Hrsg.), *Handbook of research on science education* (S. 1105–1149). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203824696-44>
- AERA, APA, & NCME (2014). *Standards for Educational and Psychological Testing*. Washington: American Educational Research Association.
- Alters, B. J., & Nelson C. E. (2002). Perspective: Teaching evolution in higher education. *Evolution: International Journal of Organic Evolution*, 56(10), 1891–901. <https://doi.org/10.1111/j.0014-3820.2002.tb00115.x>

- Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, 39(10), 952–978. <https://doi.org/10.1002/tea.10053>
- Andrews, T. M., Kalinowski, S. T., & Leonard, M. J. (2011). "Are humans evolving?" A classroom discussion to change student misconceptions regarding natural selection. *Evolution: Education and Outreach*, 4, 456–466. <https://doi.org/10.1007/s12052-011-0343-4>
- Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79–86. <https://doi.org/10.1177/0748175613513808>
- Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. *Handbook of Research on Teaching*, 4, 433–456. <https://doi.org/10.12691/education-3-4-18>
- Baalmann, W., Frerichs, V., Weitzel, H., Gropengießer, H., & Kattmann, U. (2004). Schülervorstellungen zu Prozessen der Anpassung–Ergebnisse einer Interviewstudie im Rahmen der Didaktischen Rekonstruktion. [Students' conceptions of processes of adaptation-results of an interview study in the context of didactic reconstruction]. *Zeitschrift für Didaktik der Naturwissenschaften*, 10(1), 7–28.
- Basel, N., Harms, U., Prechtel, H., Weiss, T., & Rothgangel, M. (2014). Students' arguments on the science and religion issue: the example of evolutionary theory and genesis. *Journal of Biological Education*, 48(4), 179–187.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. [Teachers' professional competence]. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133–180. <https://doi.org/10.3102/0002831209345157>
- Bautista, N. U., & Boone, W. J. (2015). Exploring the impact of TeachME™ Lab virtual classroom teaching simulation on early childhood education majors' self-efficacy

- beliefs. *Journal of Science Teacher Education*, 26(3), 237–262. <https://doi.org/10.1007/s10972-014-9418-8>
- Berti, A. E., Toneatti, L., & Rosati, V. (2010). Children's conceptions about the origin of species: A study of Italian children's conceptions with and without instruction. *Journal of the Learning Sciences*, 19(4), 506–538. <https://doi.org/10.1080/10508406.2010.508027>
- Billingsley, B., Brock, R., Taber, K. S., & Riga, F. (2016). How students view the boundaries between their science and religious education concerning the origins of life and the universe. *Science education*, 100(3), 459–482. <https://doi.org/10.1002/sce.21213>
- Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, 27(5), 415–427. <https://doi.org/10.1002/tea.3660320509>
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. (2015). Beyond dichotomies: competence viewed as a continuum. *Zeitschrift für Psychologie*, 223(3), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Bradley, E. G., & Kendall, B. (2014). A review of computer simulations in teacher education. *Journal of Educational Technology Systems*, 43(1), 3–12. <https://doi.org/10.2190/ET.43.1.b>
- Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education*, 68(4), 493–503. <https://doi.org/10.1002/sce.3730680412>
- Brunner, M., Kunter, M., Krauss, S., Baumert, J., Blum, W., Dubberke, T., ... Neubrand, M. (2006). Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? [How is the content-specific professional knowledge of mathematics teachers related to their teacher education and inservice training?]. *Zeitschrift für Erziehungswissenschaft*, 9(4), 521–544. <https://doi.org/10.1007/s11618-006-0166-1>
- Byrne, J., Grace, M., & Hanley, P. (2009). Children's anthropomorphic and anthropocentric ideas about micro-organisms: Educational research. *Journal of Biological Education*, 44(1), 37–43. <https://doi.org/10.1080/00219266.2009.9656190>
- Carlson, J., & Daehler, K. R. (2019). The refined consensus model of pedagogical content knowledge in science education. In *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (S. 77–92). Springer. https://doi.org/10.1007/978-981-13-5898-2_2

- Cheong, D. (2010). The effects of practice teaching sessions in Second Life on the change in pre-service teachers' teaching efficacy. *Computers and Education*, 55(2), 868–880. <https://doi.org/10.1016/j.compedu.2010.03.018>
- Christensen, R., Knezek, G., & Tyler-Wood, T. (2011). simSchool: An online dynamic simulator for enhancing teacher preparation. *International Journal of Learning Technology*, 6(2), 201–220. <https://doi.org/10.1504/IJLT.2011.042649>
- Clapper, T. C. (2010). Role play and simulation returning to teaching for understanding. *The Education Digest*, 75(8), 39–43.
- Dalgarno, B., Gregory, S., Knox, V., & Reiners, T. (2016). Practising teaching using virtual classroom role plays. *Australian Journal of Teacher Education*, 41(1), 126–154. <https://doi.org/10.14221/ajte.2016v41n1.8>
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, 8(1), 1–44. <https://doi.org/10.14507/epaa.v8n1.2000>
- Demastes, S. S., Settlage Jr, J., & Good, R. (1995). Students' conceptions of natural selection and its role in evolution: Cases of replication and comparison. *Journal of Research in Science Teaching*, 32(5), 535–550. <https://doi.org/10.1002/tea.3660320509>
- Depaepe, F., Torbeyns, J., Vermeersch, N., Janssens, D., Janssen, R., Kelchtermans, G., ... van Dooren, W. (2015). Teachers' content and pedagogical content knowledge on rational numbers: A comparison of prospective elementary and lower secondary school teachers. *Teaching and Teacher Education*, 47, 82–92. <https://doi.org/10.1016/j.tate.2014.12.009>
- Ferrari, M., & Chi, M. T. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education*, 20(10), 1231–1256. <https://doi.org/10.1080/0950069980201005>
- Fiedler, K., Freytag, P. & Unkelbach, C. (2007). Pseudo-contingencies in a simulated classroom. *Journal of Personality and Social Psychology*, 92, 665–677. <https://doi.org/10.1037/0022-3514.92.4.665>
- Fiedler, K., Walther, E., Freytag, P. & Plessner, H. (2002). Judgment biases in a simulated classroom – A cognitive-environmental approach. *Organizational Behavior and Human Decision Processes*, 88, 527–561. <https://doi.org/10.1006/obhd.2001.2981>
- Fischer, H. E., Borowski, A., & Tepner, O. (2012). Professional knowledge of science teachers. In *Second International Handbook of Science Education* (S. 435–448). Springer. https://doi.org/10.1007/978-1-4020-9041-7_30

- Frey, B. B. (Hrsg.). (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation*. Sage Publications. <https://doi.org/10.4135/9781506326139>
- Friedrichsen, P. J., Abdell, S., Pareja, E., Brown, P., Lankford, D., & Volkmann, M. (2009). Does teaching experience matter? Examining biology teachers' prior knowledge for teaching in an alternative certification program. *Journal of Research in Science Teaching*, 46, 357–383. <https://doi.org/10.1002/tea.20283>
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, 49(9), 1181–1210. <https://doi.org/10.1002/tea.21054>
- Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M., Girwidz, R., Obersteiner, A., ... & Seidel, T. (2018). Systematizing professional knowledge of medical doctors and teachers: development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences*, 8(4), 207–225. <https://doi.org/10.3390/educsci8040207>
- Gregory, T. R. (2009). Understanding natural selection. Essential concepts and common misconceptions. *Evolution: Education and Outreach*, 2(2), 156–175. <https://doi.org/10.1007/s12052-009-0128-1>
- Großschedl, J., Harms, U., Kleickmann, T., & Glowinski, I. (2015). Preservice biology teachers' professional knowledge: Structure and learning opportunities. *Journal of Science Teacher Education*, 26(3), 291–318.
- Großschedl, J., Mahler, D., Kleickmann, T., & Harms, U. (2014). Content-related knowledge of biology teachers from secondary schools: Structure and learning opportunities. *International Journal of Science Education*, 36(14), 2335–2366.
- Großschedl, J., Seredszus, F. & Harms, U. (2018). Angehende Biologielehrkräfte: evolutionsbezogenes Wissen und Akzeptanz der Evolutionstheorie. [Prospective biology teachers: evolution-related knowledge and acceptance of the theory of evolution.] *Zeitschrift für Didaktik der Naturwissenschaften*. 24(1), 51–70.
- Großschedl, J., Welter, V., & Harms, U. (2019). A new instrument for measuring pre-service biology teachers' pedagogical content knowledge: The PCK-IBI. *Journal of Research in Science Teaching*, 56(4), 402–439.
- Guion, R. (1977). Content validity: The source of my discontent. *Applied Psychological Measurement*, 1, 1–10. <https://doi.org/10.1177/014662167700100103>
- Gurvitch, R., & Metzler, M. W. (2009). The effects of laboratory-based and field-based practicum experience on pre-service teachers' self-efficacy. *Teaching and Teacher Education*, 25(3), 437–443. <https://doi.org/10.1016/j.tate.2008.08.006>

- Harms, U., & Reiss, M. J. (2019). The present status of evolution education. In U. Harms & M. J. Reiss (Eds.). *Evolution Education Re-considered* (S. 1–19). Basel: Springer Verlag.
- Hartmann, S., Upmeyer zu Belzen, A., Krüger, D., & Pant, H. A. (2015). Scientific reasoning in higher education: Constructing and evaluating the criterion-related validity of an assessment of preservice science teachers' competencies. *Zeitschrift für Psychologie*, 223, 47–53. <https://doi.org/10.1027/2151-2604/a000199>
- Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>
- Hattie, J., & Cooksey, R. W. (1984). Procedures for assessing the validities of tests using the "known-groups" method. *Applied Psychological Measurement*, 8(3), 295–305. <https://doi.org/10.1177/014662168400800306>
- Henze, I., & Van Driel, J. H. (2015). Toward a more comprehensive way to capture PCK in its complexity. *Re-examining Pedagogical Content Knowledge in Science Education*, 120–134.
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.
- Hixon, E., & So, H.-J. (2009). Technology's role in field experiences for preservice teacher training. *Educational Technology & Society*, 12, 294–304.
- Jamaludin, A., San Chee, Y., & Ho, C. M. L. (2009). Fostering argumentative knowledge construction through enactive role play in Second Life. *Computers & Education*, 53(2), 317–329. <https://doi.org/10.1016/j.compedu.2009.02.009>
- Jüttner, M., Boone, W., Park, S., & Neuhaus, B. J. (2013). Development and use of a test instrument to measure biology teachers' content knowledge (CK) and pedagogical content knowledge (PCK). *Educational Assessment, Evaluation and Accountability*, 25(1), 45–67. <https://doi.org/10.1007/s11092-013-9157-y>
- Jüttner, M., & Neuhaus, B. J. (2013). Validation of a paper-and-pencil test instrument measuring biology teachers' pedagogical content knowledge by using think-aloud interviews. *Journal of Education and Training Studies*, 1(2), 113–127. <https://doi.org/10.11114/jets.v1i2.126>
- Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction*, 28, 73–84.

STUDIE 2

- Kaiser, J., Südkamp, A., & Möller, J. (2017). The effects of student characteristics on teachers' judgment accuracy: Disentangling ethnicity, minority status, and achievement. *Journal of Educational Psychology*, 109(6), 871–888.
- Kalinowski, S. T., Leonard, M. J., & Taper, M. L. (2016). Development and validation of the conceptual assessment of natural selection (CANS). *CBE—Life Sciences Education*, 15(4). <https://doi.org/10.1187/cbe.15-06-0134>
- Kalthoff, H. (1995). Die Erzeugung von Wissen. Zur Fabrikation von Antworten im Schulunterricht. [The generation of knowledge. The production of answers in the classroom]. *Zeitschrift für Pädagogik*, 41(6), 925–939.
- Kampourakis, K., & Zogza, V. (2008). Students' intuitive explanations of the causes of homologies and adaptations. *Science & Education*, 17(1), 27–47. <https://doi.org/10.1007/s11191-007-9075-9>
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedm.12000>
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, 64(1), 90–106. <https://doi.org/10.1177/0022487112460398>
- Kunter, M., Baumert, J., & Blum, W. (Eds.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. [Professional competence of teachers: Results of the COACTIV research programme]. Waxmann Verlag.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project*. Springer. <https://doi.org/10.1007/978-1-4614-5149-5>
- Lawshe, C. H. (1975). A quantitative approach to content validity 1. *Personnel Psychology*, 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Lucero, M. M., Petrosino, A. J., & Delgado, C. (2017). Exploring the relationship between secondary science teachers' subject matter knowledge and knowledge of student conceptions while teaching evolution by natural selection. *Journal of Research in Science Teaching*, 54(2), 219–246. <https://doi.org/10.1002/tea.21344>
- Magnusson, S., Krajcik, J. S., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome, & N. G.

- Lederman (Hrsg.), *Examining pedagogical content knowledge* (S. 95–132). Kluwer Academic. https://doi.org/10.1007/0-306-47217-1_4
- Mahler, D., Großschedl, J., & Harms, U. (2017). Using doubly latent multilevel analysis to elucidate relationships between science teachers' professional knowledge and students' performance. *International Journal of Science Education*, 39(2), 213–237.
- Mathesius, S., Krell, M., Upmeyer zu Belzen, A., & Krüger, D. (2019). Überprüfung eines Tests zum wissenschaftlichen Denken unter Berücksichtigung des Validitätskriteriums relations-to-other-variables. [Review of a test of scientific thinking taking into account the validity criterion relations-to-other-variables]. *Zeitschrift für Pädagogik*, 4, 492–510. <https://doi.org/10.3262/ZP1904492>
- McVaugh, N. K., Birchfield, J., Lucero, M. M., & Petrosino, A. J. (2011). Evolution education: Seeing the forest for the trees and focusing our efforts on the teaching of evolution. *Evolution: Education and Outreach*, 4(2), 286. <https://doi.org/10.1007/s12052-010-0297-y>
- Nehm, R. H., Beggrow, E. P., Opfer, J. E., & Ha, M. (2012). Reasoning about natural selection: diagnosing contextual competency using the ACORNS instrument. *The American Biology Teacher*, 74(2), 92–98. <https://doi.org/10.1525/abt.2012.74.2.6>
- Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching*, 48(3), 237–256. <https://doi.org/10.1002/tea.20400>
- Nehm, R. H., Poole, T. M., Lyford, M. E., Hoskins, S. G., Carruth, L., Ewers, B. E., & Colberg, P. J. (2009). Does the segregation of evolution in biology textbooks and introductory courses reinforce students' faulty mental models of biology and evolution? *Evolution: Education and Outreach*, 2(3), 527–532. <https://doi.org/10.1007/s12052-008-0100-5>
- Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, 57(3), 263–272. <https://doi.org/10.1641/B570311>
- Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, 45(10), 1131–1160. <https://doi.org/10.1002/tea.20251>
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: knowing what students know about evolution. *Journal of Research in Science Teaching*, 49(6), 744–777. <https://doi.org/10.1002/tea.21028>

- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). *Becoming a strategic reader. Contemporary educational psychology*, 8(3), 293–316. [https://doi.org/10.1016/0361-476X\(83\)90018-8](https://doi.org/10.1016/0361-476X(83)90018-8)
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualization of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38, 261–284. <https://doi.org/10.1007/s11165-007-9049-6>
- Prinou, L., Halkia, L., & Skordoulis, C. (2008). What conceptions do Greek school students form about biological evolution? *Evolution: Education and Outreach*, 1(3), 312–317. <https://doi.org/10.1007/s12052-008-0051-x>
- Rissler, L. J., Duncan, S. I., & Caruso, N. M. (2014). The relative importance of religion and education on university students' views of evolution in the Deep South and state science standards across the United States. *Evolution: Education and Outreach*, 7(1), 1–17. <https://doi.org/10.1186/s12052-014-0024-1>
- Rozenszajn, R., & Yarden, A. (2014). Expansion of biology teachers' pedagogical content knowledge (PCK) during a long-term professional development program. *Research in Science Education*, 44(1), 189–213. <https://doi.org/10.1007/s11165-013-9378-6>
- Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal*, 50(5), 1020–1049. <https://doi.org/10.3102/0002831213477680>
- Sauvé, L., Renaud, L., Kaufman, D., & Marquis, J.-S. (2007). Distinguishing between games and simulations: A systematic review. *Journal of Educational Technology and Society*, 10, 247–256.
- Settlage Jr, J. (1994). Conceptions of natural selection: A snapshot of the sense-making process. *Journal of Research in Science Teaching*, 31(5), 449–457. <https://doi.org/10.1002/tea.3660310503>
- Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, 52(2), 170–194. <https://doi.org/10.1016/j.cogpsych.2005.10.001>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundation of the new reform. *Harvard Educational Review*, 57(1), 1–22.

- Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education and Outreach*, 1(2), 189–195. <https://doi.org/10.1007/s12052-008-0037-8>
- Sireci, S., & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, 26(1), 100–107. <https://doi.org/10.7334/psicothema2013.256>
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762.
- Südkamp, A., Möller, J., & Pohlmann, B. (2008). Der Simulierte Klassenraum: Eine experimentelle Untersuchung zur diagnostischen Kompetenz [The Simulated Classroom: An experimental investigation of diagnostic competence]. *Zeitschrift für Pädagogische Psychologie*, 22(34), 261–276.
- Tibell, L. A., & Harms, U. (2017). Biological principles and threshold concepts for understanding natural selection. *Science and Education*, 26(7), 953–973.
- Tamir, P. (1988). Subject matter and related pedagogical knowledge in teacher education. *Teaching and Teacher Education*, 4(2), 99–110. [https://doi.org/10.1016/0742-051X\(88\)90011-X](https://doi.org/10.1016/0742-051X(88)90011-X)
- Tamir, P., & Zohar, A. (1991). Anthropomorphism and teleology in reasoning about biological phenomena. *Science Education*, 75(1), 57–67. <https://doi.org/10.1002/sce.3730750106>
- Veletsianos, G., Heller, R., Overmyer, S., & Procter, M. (2010). Conversational agents in virtual worlds: Bridging disciplines. *British Journal of Educational Technology*, 41(1), 123–140. <https://10.1111/j.1467-8535.2009.01027.x>
- Yeh, Y.-C. (2004). Nurturing reflective teaching during critical-thinking instruction in a computer simulation program. *Computers & Education*, 42(2), 181–194. [https://doi.org/10.1016/S0360-1315\(03\)00071-X](https://doi.org/10.1016/S0360-1315(03)00071-X)
- Ziadie, M. A., & Andrews, T. C. (2018). Moving evolution education forward: a systematic analysis of literature to identify gaps in collective knowledge for teaching. *CBE—Life Sciences Education*, 17(1). <https://doi.org/10.1187/cbe.17-08-0190>

6 STUDIE 3: DER SIMULIERTE KLASSENRAUM BIOLOGIE – ERFASSUNG DEKLARATIVEN UND PROZEDURALEN WISSENS BEI LEHRAMTSSTUDIERENDEN DER BIOLOGIE?⁵

Zusammenfassung:

In der Lehrerprofessionalisierungsforschung werden zahlreiche Studien durchgeführt, die sich intensiv mit verschiedenen Wissensbereichen – fachliches, fachdidaktisches und pädagogisches Wissen – beschäftigen. Vernachlässigt werden häufig spezifische Wissenstypen, die nach psychologischen Ansätzen in deklaratives und prozedurales Wissen differenziert werden. Vor diesem Hintergrund untersucht die vorliegende Studie die verschiedenen Bereiche des Professionswissens angehender Biologiestudierenden (N = 51) zur Evolution und versucht gleichzeitig, erste Erkenntnisse über den Zusammenhang zwischen dem deklarativen und prozeduralen Wissen zu gewinnen. Zu diesem Zweck wird eine Kombination zweier Instrumente verwendet - ein Fragebogen zur Evolution, der das deklarative Wissen misst - und der Simulierte Klassenraum Biologie (SKR^{Bio}), der das prozedurale Wissen im Bereich der Diagnosefähigkeit erfasst. Im SKR^{Bio} können die Lehramtsstudierenden Fragen zur natürlichen Selektion an virtuelle Schüler*innen richten, woraufhin die gegebenen Antworten und die gezeigten Leistungen diagnostiziert werden müssen. Die Ergebnisse im SKR^{Bio} zeigen, dass die Lehramtsstudierenden in der Lage sind, die wissenschaftliche Korrektheit der virtuellen Schülerantworten zu beurteilen (91% Diagnoserate). Größere Schwierigkeiten werden bei der Diagnose spezifischer Fehlvorstellungskategorien innerhalb der virtuellen Schülerantworten identifiziert (59% Diagnoserate). Zwischen dem deklarativen und dem prozeduralen Wissens zeigten sich schwache bis moderate Zusammenhänge. Erwartungswidrige Ergebnisse lieferten die differenzierten Zusammenhangsanalysen zwischen deklarativem und prozeduralem Wissen innerhalb eines Wissensbereichs. Diese Befunde sind besonders relevant, da das deklarative Wissen angehender Lehrer*innen als Ausgangspunkt bzw. Voraussetzung für eine erfolgreiche Professionalisierung gilt. Um den Lehramtsstudierenden neben den schulpraktischen Phasen weitere Lerngelegenheiten zu ermöglichen, in denen prozedurales Wissen angewendet und trainiert werden kann, könnte der SKR^{Bio} eingesetzt werden.

Schlüsselwörter:

⁵ Fischer, J., Machts, N., Möller, J., & Harms, U. (major revisions, 2021). Der Simulierte Klassenraum Biologie: Erfassung deklarativen und prozeduralen Wissens bei Lehramtsstudierenden der Biologie. *Zeitschrift für die Didaktik der Naturwissenschaften*.

STUDIE 3

Professionswissen, deklaratives und prozedurales Wissen, Lehren und Lernen von Evolution, simulierte Klassenraumumgebungen

Abstract:

In teacher education research, numerous studies are conducted that intensively focus different knowledge areas – content, pedagogical-content, and pedagogical knowledge. However, specific knowledge types, which are differentiated into declarative and procedural knowledge according psychological approaches, are often neglected. Against this background, the present study investigates different areas of 51 biology pre-service teachers' professional knowledge on evolution and simultaneously tries to provide first insights into the interrelation between the respective declarative and procedural knowledge. For this purpose, a combination of two instruments is used – a questionnaire on evolution, which measures declarative knowledge – and the Simulated Classroom Biology (SCR^{Bio}), which captures procedural knowledge in the area of diagnostic ability. In the SCR^{Bio}, questions focusing on natural selection can be asked by pre-service teachers to virtual students, whereupon the given answers and the shown performance have to be diagnosed. The results in the SCR^{Bio} showed that pre-service teachers are able to assess the scientific correctness of the virtual students' answers (91% diagnostic rate). Major problems were identified in diagnosing specific misconception categories (59% diagnostic rate). Weak to moderate correlations were found between declarative and procedural knowledge. The differentiated correlation analyses between declarative and procedural knowledge within a knowledge domain provided results that were contrary to expectations. These findings are particularly relevant because declarative knowledge of pre-service teachers is considered the starting point for successful professionalization. In order to provide pre-service teachers with further learning opportunities in which procedural knowledge can be applied and trained, the SKR^{Bio} could be used.

Keyword:

Professional knowledge, Declarative and procedural knowledge, Teaching and learning evolution, Simulated classroom environments.

6.1 Einleitung

Das Lehramtsstudium bietet trotz zunehmender Praxisphasen nur wenige Möglichkeiten, das im universitären Kontext erworbene deklarative Wissen in konkreten unterrichtlichen Handlungssituationen anzuwenden und in prozedurales Wissen zu überführen. Ob und inwiefern deklaratives und prozedurales Professionswissen zusammenhängen und welche Erhebungsinstrumente sich anbieten, diese Zusammenhänge darzustellen, ist weiterhin eine offene Frage. In den vergangenen Jahren wurden zahlreiche Studien durchgeführt, welche unterschiedliche Bereiche des Professionswissens – Fachwissen (FW), fachdidaktisches Wissen (FDW) und pädagogisches Wissen (PW) (z. B. Shulman 1986) – fokussierten, sich aber kaum explizit mit dem Zusammenhang zwischen deklarativem und prozeduralem Wissen innerhalb der Professionswissensebene beschäftigten. Die vorliegende Studie hatte das Ziel, deklaratives sowie prozedurales Professionswissen zur Evolution angehender Biologielehrkräfte zu erfassen und erste Zusammenhänge aufzuzeigen. Für die Erfassung prozeduralen Wissens stellen wir hier den Simulieren Klassenraum Biologie (SKR^{Bio}) als digitales Erhebungsinstrument vor.

Das Professionswissen angehender Lehrkräfte ist der Kern professioneller Kompetenz (z. B. Borowski et al. 2010; Großschedl 2015a; Jüttner und Neuhaus 2013; Kleickmann et al. 2014; Mahler et al. 2017). Zur Konzeptualisierung des Professionswissens hat sich das aus der Bildungswissenschaft stammende COACTIV-Modell durchgesetzt, welches in FW, FDW und PW unterteilt (Baumert und Kunter 2006). Diese Wissensbereiche differenzieren sich weiter in unterschiedliche Wissensfacetten.

Neben dieser Differenzierung in Wissensbereiche und -facetten stellt die dichotome Einordnung des Professionswissens in deklaratives („Wissen, dass“) und prozedurales Wissen („Wissen, wie“) eine Möglichkeit dar, Wissen hinsichtlich der Handlungsorientiertheit zu differenzieren (z. B. Baumert und Kunter 2013; Blömeke et al. 2010; Förtsch et al. 2018; König et al. 2014; Tepner et al. 2012). Das deklarative Professionswissen wird als Sachwissen verstanden, welches dem Bewusstsein zugänglich ist und größtenteils während der ersten Phase der universitären Ausbildung erworben wird. Hingegen umfasst das prozedurale Professionswissen handlungsorientiertes Wissen (Baumert und Kunter 2013). In konkreten Handlungssituationen, wie in praktischen Übungen oder Schulpraktika, kann mittels regelmäßiger Übung und Kontextualisierung, deklaratives zu prozeduralem Wissen transformiert werden (z. B. Anderson 1982; Blömeke et al. 2010; Schneider und Stern 2010). Bei einer unvollständigen Transformation können beide Wissenstypen unverbunden nebeneinander stehen oder sich sogar widersprechen (Shulman 1986). Hier offenbart sich eine zentrale Herausforderung für die Lehrkräftebildung, wonach das im Studium erworbene deklarative Wissen in konkreten

Handlungssituationen träge bleibt und nicht in prozedurales Wissen überführt werden kann (z. B. Gruber et al. 2000).

In der Lehrkräftebildungsforschung wurden zahlreiche Large-Scale-Studien durchgeführt, die trotz der Berücksichtigung verschiedener Wissenstypen innerhalb der entwickelten Erhebungsformate überwiegend die Messung deklarativen Wissens fokussierten (*TEDS-M*: Blömeke et al. 2008; *ProwiN*: Tepner et al. 2012; *KiL*: Kleickmann et al. 2014; *COACTIV*: Krauss und Brunner 2011). Beispielsweise untersuchten Borowski et al. (2011) und Kirschner et al. (2017) mit einem Papier-Bleistift-Fragebogen das deklarative sowie prozedurale Professionswissen, wobei keine expliziten Zusammenhangsanalysen zwischen den Wissenstypen vorgenommen wurden. In den *COACTIV*-Studien wurde ebenfalls versucht, neben theoretischem Wissen (Faktenwissen) auch Facetten praktischen Wissens (Fertigkeiten) zu erfassen. Hier stellten Krauss et al. (2020) fest, dass Fragebogenformate vor allem das deklarative Wissen fokussieren, wohingegen mithilfe von Videovignetten eher situationsspezifische Fähigkeiten und Fertigkeiten erfasst werden (Blömeke et al. 2015; Lindmeier 2011, Kaiser et al. 2015). Kramer et al. (2021) konnten bspw. Zusammenhänge zwischen den Facetten des Professionswissens, erhoben über einen Fragebogen, und den gezeigten diagnostischen Aktivitäten, erfasst mit einer videobasierten Klassenraumsituation, der angehenden Biologie-Lehrkräfte nachweisen.

Die explizite Betrachtung des Zusammenhangs zwischen deklarativem und prozeduralem Wissen innerhalb der Professionswissensbereiche ist besonders relevant. Gerade die Aktivierung prozeduralen Wissens in realen Unterrichtssituation ist essentiell, um als Lehrkraft im Sinne effektiven unterrichtlichen Handelns auf die Bedürfnisse der Schüler*innen reagieren zu können. Das unterrichtliche Handeln basiert hierbei auf dem Wissen und Können des Lehrenden, welches während der universitären Ausbildung in theoretischen sowie praktischen Phasen entwickelt und im Berufskontext weiter vertieft wird. Um neben deklarativem auch prozedurales Wissen angehender Lehrkräfte zur Evolution messen und erste Zusammenhänge untersuchen zu können, bieten sich simulierte Klassenraumumgebungen an (z. B. Badiee und Kaufman 2015). In unserer Studie wurde daher eine entsprechende Umgebung weiterentwickelt und eingesetzt, die eine Klasse im evolutionsbiologischen Unterricht zum Prozess der natürlichen Selektion simuliert – der *SKR^{Bio}*. Im *SKR^{Bio}* hatten die Lehramtsstudierenden die Möglichkeit, das im akademischen Diskurs erlernte deklarative FW sowie FDW zur Evolutionstheorie und generisches PW in einer konkreten Handlungssituation anzuwenden. Hierbei ermöglichte der *SKR^{Bio}* die Messung des prozeduralen Wissens. Mit einem Fragebogen wurde das deklarative FW und FDW erfasst. Die vergleichende Betrachtung des gezeigten Wissens der Lehramtsstudierenden in beiden Instrumenten lässt eine differenzierte Betrachtung deklarativen und prozeduralen Professionswissens zur Evolution zu.

6.2 Theoretischer Hintergrund

6.2.1 Professionswissen von Biologielehrkräften

Das Professionswissen von Lehrkräften ist eine wesentliche Voraussetzung für guten Unterricht (z. B. Baumert und Kunter 2006; Borowski et al. 2010; Sadler et al. 2013, Mahler et al. 2017). Nach Shulmans (1986) Konzeptualisierung differenziert sich das Professionswissen in FW, FDW und PW. Diese Taxonomie wurde in zahlreichen Studien übernommen, um Bereiche des professionellen Wissens zu beschreiben (*ProwiN*: Fischer et al. 2012; *KiL*: Großschedel et al. 2015; Kleickmann et al. 2014; *COACTIV*: Baumert und Kunter 2013). Innerhalb der Wissensbereiche werden das FW und FDW als inhaltsbezogene Domänen beschrieben, wohingegen das PW eher als inhaltsunabhängige Domäne angesehen wird (z. B. Förtsch et al. 2018; Jüttner et al. 2013; Shulman 1986).

Das FW wird beschrieben als das Wissen von Fakten und Begriffen sowie als das konzeptuelle Verständnis dieser Inhalte (z. B. Shulman 1986). Beispielsweise basiert das evolutionsbezogene FW von Biologie-Lehrkräften auf der synthetischen Evolutionstheorie (Großschedl et al. 2015b), welche den Evolutionsprozess auf Grundlage der veränderten genetischen Zusammensetzung von Populationen beschreibt. Hierbei kann das evolutionsbezogene FW weiter in das Wissen zu den Inhaltbereichen: (1) Mikroevolution, welche unmittelbare Prozesse der genetischen Veränderung innerhalb einer Art beschreibt, (2) Makroevolution, welche Selektionsmechanismen zwischen Arten erklärt, und (3) Erklärungsmodelle, welche das Wissen zu konkurrierenden Erklärungsansätzen zur Vielfalt der Arten (z. B. nach Lamarck) zusammenfasst, differenziert werden (Großschedl et al. 2015b).

Studien haben gezeigt, dass ein elaboriertes FW für effektiven Unterricht unerlässlich ist (z. B. Baumert et al. 2010; Friedrichsen et al. 2009), es aber allein nicht ausreicht, den Herausforderungen des Unterrichts, wie zum Beispiel dem Erkennen von Fehlvorstellungen und der daraus resultierenden Wahl einer geeigneten Instruktionsstrategie, gewachsen zu sein (z. B. Baumert et al. 2010; Förtsch et al. 2018; Riese et al. 2017; Tepner et al. 2012). Hierfür ist das FDW relevant, welches vor allem durch zwei Kernfacetten – das (1) Wissen über Fehlvorstellungen und (2) das Wissen zu Instruktionsstrategien - beschrieben wird. Das Wissen über Fehlvorstellungen besteht darin, prominente Fehlvorstellungen erkennen und beschreiben zu können. Hierbei definieren Fehlvorstellungen grundlegend ungenaue Vorstellungen bzw. Ideen zu naturwissenschaftlichen Phänomenen (Yip 1998). Besonders problematisch ist die Tatsache, dass wissenschaftlich korrekte Denkweisen mit Fehlvorstellungen koexistieren können (z. B. Fosnot und Perry 2005; Shtulman und Valcarcel 2012). Nach theoretischen Ansätzen sind Fehlvorstellungen nicht pauschal lernhinderlich, sondern können auch als Anknüpfungspunkte für den Aufbau wissenschaftlicher Denkweisen im Lernprozess

dienen (Kattmann 2015). Das Wissen zu Instruktionsstrategien ist relevant, um Schüler*innen beim Erlernen wissenschaftlicher Konzepte zu unterstützen (Magnusson et al. 1999).

Im Vergleich zum FW und FDW berücksichtigt das PW übergreifende Aspekte des Unterrichts. Es wird definiert als das Wissen über Lernstrategien, effektive Unterrichts- und Klassenführung sowie Leistungsbeurteilung (z. B. Baumert und Kunter 2013; Brunner et al. 2011; Shulman 1987; König und Blömeke 2009; Voss et al. 2011). In Bezug auf die Leistungsbeurteilung ist die Beurteilungsgenauigkeit von zentraler Bedeutung, denn diese beschreibt das Verhältnis zwischen den Lehrer*innenurteilen und der tatsächlich gezeigten Schüler*innenleistungen (z. B. Südkamp et al. 2012). Ziel der Lehrkräfte muss es sein, die Schüler*innen in jeder Unterrichtssituation angemessen zu beurteilen (Schrader 2006).

Unabhängig dieser Unterscheidung des Professionswissens in Wissensbereiche wurde in psychologischen Ansätzen eine weitere Differenzierung in Wissenstypen vorgenommen. Grundlegend wird zwischen deklarativem und prozeduralem Wissen unterschieden (z. B. Anderson und Lebière 1998; Bromme 2001; de Jong und Ferguson-Hessler 1996; Fenstermacher 1994; Klieme et al. 2007; König et al. 2014). Deklaratives Wissen wird als das Wissen über Begriffe, Fakten und Prinzipien beschrieben (z. B. Baumert und Kunter 2011; Tepner et al. 2012), welches durch regelmäßige Übung und Kontextualisierung in prozedurales Wissen überführt werden kann (z. B. Anderson 1982; Blömeke et al. 2008, 2010; Klieme et al. 2007; Schneider und Stern 2010). Beispielsweise ist das Wissen über spezifische Fehlvorstellungen zur Evolution dem deklarativen Wissen zugeordnet (Förtsch et al. 2018). Geht es um die Diagnose spezifischer Fehlvorstellungen innerhalb von Schüler*innenantworten in einer Unterrichtssituation, wird prozedurales Wissen relevant. Erst wenn es in konkreten Handlungssituationen zu einer Restrukturierung und Verknüpfung des deklarativen Wissens zu prozeduralem Wissen kommt, können Lehrkräfte adäquat auf die Herausforderungen des Unterrichts reagieren (z. B. Blömeke et al. 2008).

Dieser grundlegenden Wissensstrukturierung in deklaratives und prozedurales Wissen wurde in der mathematik- sowie naturwissenschaftdidaktischen Forschung übernommen bzw. nur leicht modifiziert (z. B. Förtsch et al. 2018; Jüttner und Neuhaus 2013; Olszweski et al. 2010; Schmelzing et al. 2010b; Tepner et al., 2012). Dementsprechend wurde sich in der vorliegenden Studie der Differenzierung in deklaratives und prozedurales Wissen angeschlossen.

Diagnostisches Wissen

Lehrkräfte sind in Unterrichtssituationen kontinuierlich mit diagnostischen Aktivitäten konfrontiert (Klug et al. 2013). Hierbei ist das zugrundeliegende Professionswissen eine essentielle Voraussetzung, um diese diagnostischen Aktivitäten effektiv durchführen zu

können (z. B. Helmke et al. 2004). Diese Aktivitäten adressieren sowohl die Dokumentation des Leistungs- oder Kompetenzniveaus der Schüler*innen (z. B. Südkamp et al. 2012) als auch die Nutzung dieser diagnostischen Informationen für adaptiven Unterricht (Rogalla und Vogt 2008) und Lerninterventionen (Wollenschläger et al. 2016). In Abhängigkeit der Situation, in der eine diagnostische Aktivität benötigt wird, müssen auf unterschiedliche Bereiche des Professionswissens zurückgegriffen werden. Um beispielsweise die wissenschaftliche Korrektheit einer Schüler*inantwort einschätzen zu können, benötigen die Lehrenden entsprechendes Wissen über Fakten und Begriffe sowie ein konzeptuelles Verständnis der zu unterrichtenden Inhalte, welches dem FW zugeordnet wird (Förtsch et al. 2018, Shulman, 1986). Ist eine Fehlvorstellung die Ursache für die wissenschaftlich inkorrekte Denkweise, ist das FDW (Facette: Wissen über Schülerfehlvorstellungen; Baumert et al. 2010; Förtsch et al. 2018; Riese et al. 2017; Tepner et al. 2012) relevant.-Das PW geht über die inhaltsbezogenen Facetten hinaus und adressiert die übergeordneten Aspekte des Unterrichts (z. B. Shulman 1986, 1987). So gehört zum PW beispielsweise die Beurteilungsgenauigkeit hinsichtlich der gezeigten Leistungen der Schüler*innen unabhängig vom konkreten Fachinhalt (Voss und Kunter 2011). Die Beurteilungsgenauigkeit kann durch drei Komponenten gemessen werden: (1) Die Rangkomponente, die zeigt, ob die Lehrenden in der Lage sind, die charakteristische Ausprägung der Schüler*innen untereinander richtig zu beurteilen. (2) Die Niveauelemente gibt die Fähigkeit der Lehrenden an, das durchschnittliche Niveau eines Schüler*innenmerkmals in einer Klasse richtig einzuschätzen und (3) die Differenzierungskomponente, als Verhältnis der Varianz der realen Leistungen zur Varianz der eingeschätzten Leistungen (Südkamp et al. 2008).

Die vorliegende Studie fokussierte die Untersuchung des diagnostischen Wissens in den drei Wissensbereichen (FW, FDW, PW), wobei der Fokus ausschließlich auf der Beurteilung lag, ohne hierbei weitere Facetten des professionellen Wissens, wie bspw. Aspekte adaptiven Unterrichts oder explizite Instruktionsstrategien mit einzubeziehen.

6.2.2 Lehren und Lernen der biologischen Evolution

Die Evolutionstheorie gilt als übergeordnetes Organisationsprinzip der Biologie, welches einzelne Fakten zu einem umfassenden Erklärungsrahmen verbindet (Nationale Akademie der Wissenschaft Leopoldina 2017; National Research Council [NRC] 2012). Evolutionäre Prozesse sind die Grundlage aller biologischen Themen und eines konzeptionellen Verständnisses der Biologie (z. B. Anderson et al. 2002; Basel et al. 2014; Bishop und Anderson 1990; Furtak 2012; Opfer et al. 2012; Zabel und Gropengießer 2011). Folglich ist die Evolutionstheorie in vielen Ländern zentraler Bestandteil zahlreicher Lehrpläne im naturwissenschaftlichen Unterricht (z. B. Deutschland: Ständige Konferenz der

Kultusminister der Länder in der Bundesrepublik Deutschland [KMK] 2005; USA: Next Generation Science Standards [NGSS] 2013).

Trotz der starken Fokussierung in den vergangenen Jahren belegen zahlreiche Studien erhebliche Schwierigkeiten beim Lehren und Lernen der Evolutionstheorie (z. B. BouJaoude et al. 2011; Gregory 2009; Großschedl et al. 2014; Lucero et al. 2019; Nehm und Reilly 2007; Nehm und Schonfeld 2007). Im Kontext des naturwissenschaftlichen Unterrichts bleiben nicht-wissenschaftliche Vorstellungen, im Folgenden Fehlvorstellungen⁶ genannt, von Schülern häufig unberührt, wodurch es weniger Gelegenheiten gibt, den wissenschaftlichen Erklärungsansatz zu thematisieren (z. B. Kampourakis und Zogza 2008; McVaugh et al. 2011; Opfer et al. 2012). Fehlvorstellungen werden von Schülern als Ideen und Gedanken beschrieben, welche mit den wissenschaftlichen Erkenntnissen unvereinbar sind (Yip 1998). Diese Fehlvorstellungen zur Evolution sind sehr vielfältig und berühren unterschiedliche evolutionäre Mechanismen. So wurden von Graf und Hamdorf (2011) über 80 Fehlvorstellungskategorien identifiziert, welche vor allem den Evolutionsbegriff, die zeitliche Dimension von Evolution und die Wissenschaftlichkeit der Evolutionstheorie betreffen. Besonders zum Prozess der natürlichen Selektion wurden zahlreiche Fehlvorstellungen bei Schülern identifiziert und definiert, welche sich vor allem auf die Erklärung der Anpasstheit von Individuen und Populationen an ihre Umwelt beziehen (z. B. Baalman et al. 2004; Bishop und Anderson 1990; Ferrari und Chi 1998; Lucero et al. 2019).

6.2.3 Fehlvorstellungen zum Prozess der natürlichen Selektion

Die natürliche Selektion ist der zentrale Prozess des evolutionären Wandels, welcher zu adaptiven Merkmalsveränderungen bei Organismen führt. Aus der Forschung ist bekannt, dass der Prozess der natürlichen Selektion eines der am schwierigsten zu vermittelnden Themen in der Biologie ist (z. B. Bishop und Anderson, 1990; Nehm und Reilly, 2007). Verantwortlich sind zahlreiche Fehlvorstellungen seitens der Lernenden, welche von den Lehrkräften erkannt und überwunden werden müssen (Gregory 2009). Überwiegend beruhen diese Fehlvorstellungen auf der Annahme, dass nicht Faktoren wie Mutation und Rekombination Ursprung der natürlichen Selektion sind, sondern sich Lebewesen aktiv und wissentlich neuen Umweltbedingungen anpassen (z. B. Settlage 1994; Nehm und Schonfeld 2008). In den vergangenen Jahren versuchten zahlreiche Forschungsarbeiten spezifische Fehlvorstellungen zu identifizieren und Kategorisierungen einzuführen. Es kristallisierte sich heraus, dass vor allem anthropomorphe, teleologische und lamarckistische Fehlvorstellungen von den Lernenden genutzt werden, um den Prozess der natürlichen Selektion zu erklären (z. B. Bishop und Anderson 1990; Bizzo 1994; Demastes

⁶ Nach Graf und Hamdorf (2011, S. 26)

et al. 1995; Hammann und Nehm 2020; Jensen und Finley 1996; Kampourakis und Zogza 2008; Nehm et al. 2009).

Die anthropomorphe Fehlvorstellung ist dadurch gekennzeichnet, dass menschliche Eigenschaften, wie zum Beispiel das Denken oder Fühlen, auf nichtmenschliche Organismen übertragen werden (z. B. Byrne et al. 2009). Die Veränderung eines Merkmals ist das Ergebnis von absichtlichen und zielgerichteten Verhalten oder Handlungen durch das Individuum selbst, um sich aktiv an veränderte Umweltbedingungen anzupassen (z. B. Baalman et al. 2004; Demastes et al. 1995; Gregory 2009; Kallery und Psillos 2004; Sinatra et al. 2008). Auch in teleologischen Erklärungen werden Veränderungen mit einer Ziel- bzw. Zweckgerichtetheit begründet, welche durch einen Schöpfer oder das Lebewesen selbst gesteuert werden. Merkmale entstehen oder gehen verloren, weil diese zu einem Überlebensvorteil für das Individuum führen (z. B. Andrews et al. 2011; Nehm et al. 2009; Sinatra et al. 2008). Teleologische Erklärungen folgen immer einem „Start-Ziel-Schema“ mit einem manifesten Endergebnis (z. B. Sinatra et al. 2008; Stover und Mabry 2007). Nach lamarckistischen Fehlvorstellungen entstehen Merkmalsveränderungen durch die aktive Anpassung der Individuen während der Nutzung oder Nichtnutzung von Organen bzw. Fähigkeiten. (z. B. Andrews et al. 2011; Kalinowski et al. 2016; Kampourakis und Zogza 2008; Nehm et al. 2009).

Diese ziel- und zweckgerichteten Erklärungen basieren auf der natürlichen Tendenz des Menschen, Naturphänomene kausal erklären zu wollen (z. B. Gregory 2009; Olander 2012). Dennoch ist nicht das Ziel oder der Zweck der bestimmende Faktor für die Entwicklung eines Merkmals, sondern evolutionsbiologische Mechanismen wie Variabilität, Selektion und Vererbung (z. B. Tibell und Harms 2017). Das konzeptuelle biologische Wissen, welches zur Erklärung evolutionsbiologischer Prozesse notwendig ist, sollte durch den Biologieunterricht vermittelt werden (z. B. Gregory, 2009). Hierfür ist ein elaboriertes Wissen der Lehrkräfte bezüglich prominenter Fehlvorstellungen zum Prozess der natürlichen Selektion notwendig, um im Unterricht geäußerte Fehlvorstellungen zur Evolution zu erkennen und mögliche Anschlusshandlungen zur Initiierung wissenschaftlich korrekter Denkweisen umzusetzen.

6.3 Fragestellung und Hypothesen

Das deklarative Wissen, das angehende Lehrkräfte während ihres Studiums erlernen, bleibt in konkreten Unterrichtssituationen oft träge und wird nicht abgerufen (z. B. Blömeke et al. 2015). Dies scheint eine der größten Herausforderungen in der naturwissenschaftlichen Lehrkräftebildung zu sein. Dementsprechend sollte Lehramtsstudierenden die Möglichkeit gegeben werden, ihr deklaratives Wissen in konkreten Handlungssituationen anzuwenden

und prozedurales Wissen zu entwickeln (Santagata und Guarino 2011). Hierfür müsste vorerst in der Forschung untersucht werden, inwiefern das deklarative und prozedurale Wissen bei angehenden Biologielehr*innen zur Evolution zusammenhängt. Eine Voraussetzung sind Instrumente zur Erfassung deklarativen und prozeduralen Wissens, um einen tieferen Einblick in deren Zusammenhang zu gewähren.

Hieraus ergibt sich die zentrale Forschungsfrage, inwieweit das deklarative und prozedurale Wissen der Lehramtsstudierenden zusammenhängen. Basierend auf theoretischen Vorkenntnissen aus der Wissenspsychologie (z. B. Anderson 2001; Edelmann 2000; Steiner 2001) sowie empirischen Befunden (Rittle-Johnson et al. 2001; Perry, 1991), wonach deklaratives und prozedurales Wissen miteinander interagiert, gehen wir davon aus, dass Zusammenhänge zwischen dem deklarativen Wissen im Fragebogen und dem prozeduralen Wissen im SKR^{Bio} bestehen (Hypothese 1). Für die differenziertere Betrachtung spezifischer Wissensbereiche (FW, FDW und PW) erwarten wir darüber hinaus vergleichsweise höhere Zusammenhänge zwischen dem deklarativen FW im Fragebogen und dem prozeduralen FW im SKR^{Bio} (Hypothese 2) sowie zwischen dem deklarativen FDW im Fragebogen und dem prozeduralen FDW im SKR^{Bio} (Hypothese 3).

6.4 Methodisches Vorgehen

6.4.1 Stichprobe und Ablauf

Die Studie wurde an einer norddeutschen Universität durchgeführt. Es nahmen $N = 51$ Lehramtsstudierende (83% weiblich) im Master of Education für das gymnasiale Lehramt mit dem Unterrichtsfach Biologie teil. Die Teilnehmenden waren im Mittel $M = 25,7$ ($SD = 2,97$) Jahre alt.

6.4.3 Testinstrumente

ProWiE - Fragebogen - Messung des deklarativen Professionswissens

Das deklarative FW und FDW wurde mit einem Fragebogen erfasst (Großschedl et al. 2015b: Testinstrument zur Erfassung des evolutionsbezogenen Professionswissens von Lehramtsstudierenden; ProWiE). 28 geschlossene Fragen des Fragebogens bezogen sich auf das evolutionsbezogene FW. 21 Fragen (4 im geschlossenen, 17 im offenen Format) adressierten das evolutionsbezogene FDW.

Wir verwendeten die aus der IRT Skalierung des Fragebogens erhaltenen Personenfähigkeitsschätzer (Weighted Likelihood Estimates) als Indikator für das deklarative Wissen, wobei der Gesamttest – FW und FDW – eine akzeptable Reliabilität zeigte ($rel.EAP = 0,64$). Nach Großschedl et al. (2015b) differenzierten wir das gezeigte

Wissen im Fragebogen weiter in deklaratives FW und deklaratives FDW, welche auf Basis der IRT-Analysen mäßige bis akzeptable Reliabilitäten zeigten (FW: $rel.EAP = 0,63$; FDW: $rel.EAP = 0,54$).

Der Simulierte Klassenraum Biologie – Messung des prozeduralen Professionswissens

Der SKR^{Bio} ist eine simulierte Klassenraumumgebung, welche es den Lehramtsstudierenden ermöglicht, in die Rolle der Lehrenden zu schlüpfen und virtuell zu unterrichten (Südkamp et al. 2008). Im Fokus geht es um die Beurteilung von Unterrichtsleistungen der virtuellen Schüler*innen, wobei der SKR^{Bio} keine komplexen Unterrichtssituationen darstellen kann. Während dieser digitalen Unterrichtsstunde können Fragen an virtuelle Schüler*innen gerichtet werden, welche durch ein Profilbild und dem dazugehörigen Namen dargestellt werden. Die virtuellen Schüler*innen antworten immer bezüglich ihres voreingestellten Fähigkeitsprofils. Es betrifft bspw. die Wahrscheinlichkeit eine richtige Antwort zu geben oder die Häufigkeit, sich auf eine gestellte Frage zu melden. Auf Basis der Qualität der gegebenen Antworten können die Leistungen der virtuellen Schüler*innen eingeschätzt werden. Die Übereinstimmung zwischen den Einschätzungen der Lehramtsstudierenden und der tatsächlich erbrachten Leistung der virtuellen Schüler*innen ergibt ein Maß für die Beurteilungsgenauigkeit.

Um neben PW auch FW und FDW angehender Lehrkräfte erfassen zu können, wurde der SKR (Fiedler et al. 2002, 2007; Südkamp et al. 2008) zum SKR^{Bio} weiterentwickelt. Hierfür haben wir eine virtuelle Unterrichtsstunde zur biologischen Evolution mit 27 evolutionsbiologischen Fragestellungen sowie dazugehörigen Schüler*innenantworten ($N = 476$) erstellt (siehe Online Material). Biologiedidaktische Studien zeigen, dass das Wissen zur natürlichen Selektion mit der jeweils betroffenen Organismengruppe divergiert (Opfer et al. 2012, Großschedl et al. 2018). Darüber hinaus gibt es Erkenntnisse darüber, dass die Merkmalsausprägung (d.h. Verlust vs. Erwerb) einen Einfluss auf das Verständnis des Evolutionsprozesses haben kann (Nehm et al. 2012). Dementsprechend adressieren die evolutionsbiologischen Fragestellungen im SKR^{Bio} alle Organismengruppen (Bakterien, Pflanzen, Tiere und Mensch) und verschiedenen Merkmalsausprägungen. Die entwickelten Schüler*innenantworten basieren u.a. auf authentischen Schüler*innenantworten einer Interviewstudie (Baalman et al. 2004) und äußern entweder eine wissenschaftliche Denkweise oder eine spezifische Fehlvorstellung (anthropomorph, lamarckistisch, teleologisch).

In der vorliegenden Studie wurde die virtuelle Klasse in drei Gruppen mit je drei virtuellen Schüler*innen (insgesamt 9 virtuelle Schüler*innen, s. Abb. 1) aufgeteilt, wobei die Wahrscheinlichkeit wissenschaftlich korrekter Antworten zwischen den drei Schüler*innengruppen bei 20, 50 und 80 Prozent lag. Außerdem wurde jedem/jeder virtuellen Schüler*in eine spezifische Fehlvorstellungskategorie hinterlegt. Das heißt,

antworteten die Schülerin*innen auf die gestellte Frage fehlerhaft, so geschah dies immer auf Basis der für diese/n Schüler*in voreingestellte Fehlvorstellungskategorie. Innerhalb des SKR^{Bio} wurde jede mögliche Wahrscheinlichkeit eine wissenschaftliche korrekte Antwort zu geben (d.h., 20, 50 oder 80 Prozent) mit jeweils einer spezifischen Fehlvorstellungskategorie (d.h. anthropomorph, teleologisch, lamarckistisch: dominante Fehlvorstellung) kombiniert.

Die Aufgabe der Lehramtsstudierenden war es, während der 30-minütigen digitalen Unterrichtsstunde Fragen aus einem Fragenmenü auszuwählen und die darauf gegebenen virtuellen Schüler*innenantworten hinsichtlich ihrer wissenschaftlichen Korrektheit bzw. der jeweils artikulierten Fehlvorstellungskategorie zu diagnostizieren. Nach Beendigung der virtuellen Unterrichtsstunde mussten die individuellen Leistungen der virtuellen Schüler*innen, d.h. Anteil richtiger Antworten während der gesamten Unterrichtsstunde, beurteilt sowie die am häufigsten verwendete (dominante) Fehlvorstellung diagnostiziert werden.

The screenshot displays the SKR Bio interface during a virtual lesson. At the top, it shows 'Klassenraum' and a token 'xEcVYW'. Below this is a grid of nine virtual student avatars, each with a name tag: Nidger, Michelle, Tim, Lukas, Laura, Lea, Finn, Jan, and Sarah. To the right of the avatars, a text box explains the 'Virtuelle Klasse' concept. Below the avatars is a 'Fragenmenü' (Question Menu) with four columns: 'Kategorien', 'Unterkategorien', 'Fragen', and 'Frage an Klasse gerichtet'. The 'Kategorien' column has 'Unterrichtsfragen' selected. The 'Unterkategorien' column lists 'Zoologie', 'Botanik', 'Humanbiologie', and 'Bakteriologie'. The 'Fragen' column lists various topics like 'Nachtlerzen - Giftstoffe', 'Teufelszwirn - Pflanzenorgane', 'Tabakpflanze - Abwehrmechanismen', 'Spaltöffnungen bei Landpflanzen', 'Gauklerblumen - Blütenmorphologie', 'Rotes Straußgras - Kupferresistenz', and 'Weißklee - Cyanidproduktion'. The 'Frage an Klasse gerichtet' column contains two questions about 'Rotes Straußgras' and 'Weißklee'. A note at the bottom explains that the 'Fragen - Pflanzen' section contains seven questions selected by the students.

Kategorien	Unterkategorien	Fragen	Frage an Klasse gerichtet
Unterrichtsfragen	Zoologie	Nachtlerzen - Giftstoffe	Das Rote Straußgras ist ein weit verbreitetes Gras, welches auch schon in der Nähe von Kupferminen gefunden wurde. An diesen Stellen sind die Böden stark kupferbelastet und wirken auf viele Pflanzen toxisch. Untersuchungen zeigten, dass das rote Straußgras in diesen Bereichen eine hohe Kupfertoleranz aufweist, welche bei Artgenossen außerhalb von kupferbelasteten Böden nicht festgestellt werden konnte. Wie ist die Kupfertoleranz unter Berücksichtigung evolutionsbiologischer Prinzipien zu erklären?
	Botanik	Teufelszwirn - Pflanzenorgane	
	Humanbiologie	Tabakpflanze - Abwehrmechanismen	Frage an die Klasse : Es wird die vollständige Frage angezeigt, auf welche die virtuellen Schüler antworten sollen.
	Bakteriologie	Spaltöffnungen bei Landpflanzen	
		Gauklerblumen - Blütenmorphologie	
		Rotes Straußgras - Kupferresistenz	
		Weißklee - Cyanidproduktion	

Fragen - Pflanzen: Innerhalb der Pflanzen konnten sieben Fragen von den Lehramtsstudierenden ausgewählt werden (hier: Frage zur Kupferresistenz zum roten Straußgras).

Abbildung 6.1 Darstellung des SKR^{Bio} während der Fragenauswahl durch die Lehramtsstudierenden.

6.4.4 Operationalisierung von deklarativem und prozeduralem Wissen

Das gezeigte Wissen im Fragebogen wurde als deklaratives Wissen operationalisiert, da die Lehramtsstudierenden zur Beantwortung auf Wissen zu Fakten, Begriffen und Prinzipien zurückgreifen mussten (z. B. Tepner et al. 2012).

Das gezeigte FW, FDW und PW der Lehramtsstudierenden im SKR^{Bio} wurde als prozedurales Wissen operationalisiert, da sich die entsprechenden Wissensbereiche hier in einer konkreten Handlungssituation artikulierten (z. B. Jüttner und Neuhaus 2013; Tepner et al. 2012), indem die Lehramtsstudierenden nach Abgabe einer Antwort sowie nach Abschluss der digitalen Unterrichtsstunde die virtuellen Schüler*innen direkt beurteilen mussten.

Für die Erfassung des prozeduralen Wissens im SKR^{Bio} wurden vier Variablen operationalisiert:

1. Der Anteil korrekter Diagnosen wissenschaftlich korrekter virtueller Schüler*innenantworten (Wissen über die Evolutionstheorie; betrifft FW)
2. Der Anteil korrekter Diagnosen spezifischer Fehlvorstellungen in den virtuellen Schüler*innenantworten (Facette: Wissen über Fehlvorstellungen; betrifft FDW)
3. Die Beurteilungsgenauigkeit am Ende der Unterrichtsstunde (Facette: pädagogisches Beurteilungswissen; betrifft PW)
4. Der Anteil korrekter Diagnosen der dominanten Fehlvorstellungen bei einzelnen virtuellen Schüler*innen am Ende der Unterrichtseinheit (Facette: Wissen über Fehlvorstellungen; betrifft FDW)

Die Anteile korrekter Diagnosen (1) wissenschaftlich korrekter und fehlerhafter Antworten, (2) spezifischer Fehlvorstellungskategorien, (3) sowie dominanter Fehlvorstellungen im SKR^{Bio} werden als relative Häufigkeiten angegeben.

Die Beurteilungsgenauigkeit der Lehramtsstudierenden im SKR^{Bio} hinsichtlich der individuellen Leistungen der virtuellen Schüler*innen am Ende der Unterrichtssequenz wurde über drei Beurteilungskomponenten (Rang-, Niveau- und Differenzierungskomponente) ermittelt (Schrader und Helmke 2001; Spinath 2005).

6.4.5 Ablauf

Die Untersuchung dauerte etwa 90 – 120 Minuten und beinhaltete zwei Teile. Zuerst wurden die Lehramtsstudierenden gebeten, in etwa 45 – 60 Minuten einen Fragebogen zur Evolution zu beantworten. Anschließend arbeiteten die Lehramtsstudierenden mit dem SKR^{Bio} (ca. 45 – 60 Minuten).

6.4.6 Datenauswertung

Erfassung deklarativen Wissens

Die Antworten im Fragebogen wurden mit dem Kodiermanual der ProWiE-Studie ausgewertet (Großschedl et al. 2015b). Offene Antwortformate ($N = 17$) wurden von einem weiteren unabhängigen Rater codiert, um eine adäquate Punktzuweisung zu gewährleisten. Die aus der IRT Skalierung des Fragebogens erhaltenen Personenfähigkeitsschätzer (deklaratives Wissen) wurden genutzt, um diese mit den Operationalisierungen für prozedurales Wissen im SKR^{Bio} zu vergleichen.

Erfassung prozeduralen Wissens

Die diagnostischen Leistungen im SKR^{Bio} der Lehramtsstudierenden wurden auf Individualebene anhand absoluter Werte korrekter bzw. inkorrektur Diagnosen der virtuellen Schüler*innenantworten errechnet.

Zusammenhang zwischen dem deklarativen und prozeduralen Wissen

Die aus der IRT Skalierung des Fragebogens erhaltenen Personenfähigkeitsschätzer (deklaratives Wissen) wurden genutzt, um diese mit den Operationalisierungen für prozedurales Wissen im SKR^{Bio} zu vergleichen. Hierbei wurden Zusammenhänge mittels Korrelationsanalysen (Produkt-Moment-Korrelationen) ermittelt, wobei die Korrelationen gegen 0 sowie gegeneinander getestet wurden.

Um die Beurteilungsgenauigkeit zu ermitteln, wurden die eingeschätzten Leistungen der virtuellen Schüler*innen durch die Lehramtsstudierenden mit den tatsächlich gezeigten Leistungen der virtuellen Schüler*innen im SKR^{Bio} verglichen. Zur Untersuchung des Zusammenhangs zwischen deklarativem Wissen und der Beurteilungsgenauigkeit im SKR^{Bio} (prozedurales Wissen) wurde nur die Rangkomponente berücksichtigt, da die Niveau- sowie die Differenzierungskomponente sich in unserer Studie nicht linear verhalten und damit eine Korrelation nicht berechnet werden konnte.

6.5 Ergebnisse

6.5.1 Deklaratives Wissen

Im Fragebogen erreichten die Lehramtsstudierenden durchschnittlich $M = 35,69$ von 67 Punkten (Min. = 22 Punkte; Max. = 47 Punkte; $SD = 5,89$). Im Mittel erreichten die Lehramtsstudierenden im fachwissenschaftlichen Teil des Fragebogens $M = 19,67$ von 32 möglichen Punkten (Min. = 11 Punkte; Max. = 28 Punkte; $SD = 3,69$). Im fachdidaktischen Teil des Fragebogens erreichten die Studierenden im Mittel $M = 16,02$ von 35 möglichen Punkten (Min. = 7 Punkte; Max. = 24 Punkte; $SD = 3,65$). Mittels Produkt-Moment

Korrelationen der gewichteten Personenfähigkeitsschätzer innerhalb der Stichprobe konnte zwischen *FW* und *FDW* im Fragebogen ein schwacher Zusammenhang ermittelt werden ($r = 0,29$; $p < 0,05$).

6.5.2 Prozedurales Wissen

Insgesamt wurden $N = 2988$ Fragen durch die Lehramtsstudierenden im SKR^{Bio} gestellt, auf welche die virtuellen Schüler*innen zu 49,2% ($n = 1471$) wissenschaftlich korrekt und zu 50,8% ($n = 1517$) fehlerhaft antworteten. Auf Basis gegebener Antworten der virtuellen Schüler*innen waren die Lehramtsstudierenden in der Lage, 91,8 % ($n = 1350$) der wissenschaftlich korrekten Antworten und 90,1 % ($n = 1367$) der fehlerhaften Antworten zu erkennen.

Die Lehramtsstudierenden erkannten im Mittel zu 58,7 % ($n = 891$) die spezifische Fehlvorstellungskategorie. Am häufigsten konnte die lamarckistische Fehlvorstellung identifiziert werden (68,6 %; $n = 382$), gefolgt von der teleologischen Fehlvorstellung 57,5% ($n = 279$). Die anthropomorphe Fehlvorstellung wurde mit 50,2 % ($n = 230$) am seltensten erkannt.

Über alle Lehramtsstudierenden hinweg ergab sich eine mittlere Rangkomponente von $M_R = 0,39$ (Min. = -0,47; Max. = 0,89), eine mittlere Niveauelemente von $M_N = -1,17$ (Min. = -38,66; Max. = 16,11) und eine mittlere Differenzierungskomponente von $M_D = 0,88$ (Min. = 0,23; Max. = 1,65). Die Ergebnisse der Beurteilungsgenauigkeit zeigen eine moderate Übereinstimmung zwischen der Beurteilung der Lehramtsstudierenden bezüglich der Leistungen der virtuellen Schüler*innen im SKR^{Bio} (Rangkomponente), wobei die Leistungen minimal unterschätzt wurden (Niveauelemente). Innerhalb der virtuellen Klasse wurde die Varianz der virtuellen Schüler*innenleistungen im Mittel leicht unterschätzt (Differenzierungskomponente).

6.5.3 Zusammenhang deklarativen und prozeduralen Wissens

Zwischen dem deklarativen Wissen im ProWiE-Fragebogen und den Operationalisierungen prozeduralen Wissens im SKR^{Bio} zeigten sich entsprechend Hypothese 1 schwache bis moderate Zusammenhänge (s. Tab. 1). Die Anteile korrekter Diagnosen wissenschaftlich korrekter virtueller Schüler*innenantworten ($Wiss_{Korr}$: $r = 0,38$; $p = 0,00$), spezifischer Fehlvorstellungen in virtuellen Schüler*innenantworten (FV_{Hit} : $r = 0,45$; $p = 0,00$) sowie der dominanten Fehlvorstellungen bei einzelnen virtuellen Schüler*innen am Ende der Unterrichtsstunde (FV_{Dom} : $r = 0,25$, $p = 0,04$) korrelierten statistisch signifikant mit dem gezeigten deklarativen Wissen. Eine genauere Betrachtung der Diagnosen spezifischer Fehlvorstellungen zeigte, dass der Anteil korrekter Diagnosen bei Antworten mit anthropomorphen (FV_{Anth} : $r = .23$, $p = .04$) und teleologischen

Fehlvorstellungen ($FV_{Tel}: r = 0,45, p = 0,00$) statistisch signifikant mit dem deklarativen Wissen korreliert, bei lamarckistischen Fehlvorstellungen scheint die Diagnose hingegen nicht mit deklarativem Wissen zusammenzuhängen ($FV_{Lam}: r = 0,06, p = 0,67$). Die Korrelation mit der Beurteilungsgenauigkeit (Rangkomponente) war statistisch knapp nicht signifikant ($RK: r = 0,19; p = 0,09$).

Bei einer detaillierten Betrachtung beider Wissensbereiche (FW und FDW) innerhalb des Fragebogens (deklaratives Wissen) zeigten sich augenscheinlich Unterschiede in der Größe der Zusammenhänge mit den Operationalisierungen prozeduralen Wissens. Die Anteile korrekter Diagnosen wissenschaftlich korrekter virtueller Schüler*innenantworten (prozedurales Wissen) zeigten größere Zusammenhänge mit dem deklarativen FDW ($Wiss_{Korr}: r = 0,38; p = 0,00$) als mit dem deklarativen FW ($Wiss_{Korr}: r = 0,24; p = 0,09$), womit Hypothese 2 zurückgewiesen wird. Weiterhin konnten in Widerspruch zu Hypothese 3 höhere Zusammenhänge zwischen den Anteilen korrekter Diagnosen spezifischer Fehlvorstellungen (prozedurales Wissen) und dem deklarativen FW beobachtet werden ($FV_{Anth}: r = 0,34, p = 0,01; FV_{Tel}: r = 0,41, p = 0,00$) als zwischen den Anteilen korrekter Diagnosen spezifischer Fehlvorstellungen (prozedurales Wissen) und dem deklarativen FDW ($FV_{Anth}: r = 0,03, p = 0,82; FV_{Tel}: r = 0,28, p = 0,04$). Jedoch waren keine der Zusammenhangsunterschiede zwischen dem deklarativen FW und dem deklarativen FDW statistisch signifikant (siehe Tabelle 1).

STUDIE 3

Tabelle 6.1 Produkt-Moment-Korrelationen zwischen dem gezeigten deklarativen Wissen im Fragebogen und dem gezeigten prozeduralen Wissen im SKR^{Bio} der Lehramtsstudierenden.

Prozedurales Wissen	Deklaratives Wissen		
	Fachwissen ^a	Fachdidaktisches Wissen ^b	Fachliches – und fachdidaktisches Wissen ^c
Wiss _{Korr}	.24	.38**	.38**
FV _{Anth}	.34*	.03	.23*
FV _{Tel}	.41**	.28*	.45**
FV _{Lam}	-.00	.06	.06
FV _{Hit}	.35*	.21	.37**
FV _{Dom}	.26	.15	.25*
RK	.17	.14	.19

Anmerkungen. ^a*n* = 28 fachliche Fragen. ^b*n* = 21 fachdidaktische Fragen. ^c*N* = 49 fachliche und fachdidaktische Fragen (gesamter Fragebogen). WissKorr = Anteil korrekter Diagnosen wissenschaftlich korrekter Antworten. FV_{Anth} = Anteil korrekter Diagnosen anthropomorpher Fehlvorstellungen. FV_{Tel} = Anteil korrekter Diagnosen teleologischer Fehlvorstellungen. FV_{Lam} = Anteil korrekter Diagnosen lamarckistischer Fehlvorstellungen. FV_{Hit} = Anteil korrekter Diagnosen spezifischer Fehlvorstellungskategorien. FV_{Dom} = Anteil korrekter Diagnosen dominanter Fehlvorstellungen. RK = Beurteilungsgenauigkeit auf Basis der Rangkomponente.

* $p < .05$. ** $p < .01$.

6.6 Diskussion

Die im Artikel vorgestellte Studie zum deklarativen und prozeduralen evolutionsbezogenen Wissen angehender Biologielehrkräfte versucht mit einer Kombination zweier Erhebungsinstrumente: (1) dem ProwiE-Fragebogen und dem (2) SKR^{Bio}, beide Wissenstypen separat zu erfassen, um im Anschluss erste Zusammenhänge darstellen zu können. Grundlage für die Interpretation des gezeigten deklarativen und prozeduralen Wissens der Lehramtsstudierenden ist die Annahme, dass die eingesetzten Messinstrumente in der Lage sind, die jeweils fokussierten Wissenstypen adäquat abzubilden. Da die Konzeptualisierungen von Wissenstypen in der kognitions- und bildungspsychologischen sowie naturwissenschaftsdidaktischen Forschung zum Teil stark divergieren, stellt sich die Operationalisierung von deklarativen und prozeduralen Wissen als herausfordernd dar.

6.6.1 Gezeigtes prozedurales Wissen im SKR^{Bio}

Innerhalb des SKR^{Bio} konnte das prozedurale Professionswissen (FW, FDW, PW) der Biologie-Lehramtsstudierenden erfasst werden. Zur Beurteilung der Wissenschaftlichkeit einer virtuellen Schüler*innenantwort mussten die Lehramtsstudierenden auf ihr FW zur synthetischen Evolutionstheorie zurückgreifen. Die hohen Anteile korrekter Diagnosen wissenschaftlich korrekter und fehlerhafter virtueller Schüler*innenantworten lassen vermuten, dass der Prozess der natürlichen Selektion von den Lehramtsstudierenden grundsätzlich verstanden wurde und das wissenschaftliche sowie unwissenschaftliche Aspekte einer Schüler*inantwort zuverlässig erkannt wurden. Obwohl das FW eine zentrale Bedeutung für die Qualität von Unterricht hat (z. B. Baumert et al. 2010; Friedrichsen et al. 2009), ist es ebenso Konsens, dass FW um FDW ergänzt werden muss, um erfolgreichen Unterricht durchzuführen und den Lernerfolg der Schüler*innen zu garantieren (z. B. Abell 2007; Baumert et al. 2010). Eine relevante Facette des FDW ist hierbei das Wissen über Fehlvorstellungen (z. B. Förtsch et al. 2018; Schmelzing et al. 2013), welches im SKR^{Bio} angewendet werden musste, um die virtuellen Schüler*innenantworten hinsichtlich spezifischer Fehlvorstellungskategorien innerhalb der Unterrichtsstunde zu diagnostizieren. Brumby (1984) zeigte in seinen Studien, dass etwa zwei Drittel der Biologiestudierenden Fehlvorstellungen während der Erklärung des Prozesses der natürlichen Selektion äußerten. Besonders prominent waren auch hier teleologische und lamarckistische Fehlvorstellungen. In Untersuchungen von Nehm und Schonfeld (2008) konnten diese Ergebnisse repliziert werden. Teleologische, anthropomorphe und lamarckistische Denkweisen basieren auf der Tendenz des Menschen, natürliche Phänomene kausal erklären zu wollen (z. B. Olander 2012). Hier liegt die Vermutung nahe, dass Fehlvorstellungen basierend auf einer kausalen Erklärung eher von Lehramtsstudierenden akzeptiert und somit seltener diagnostiziert werden (z. B. Gresch und Martens 2019; Kallery und Psillos 2004; Kampourakis und Zogza 2008; Sinatra et al. 2008).

Bezüglich des PW zeigte sich, dass die Lehramtsstudierenden am Ende der virtuellen Unterrichtsstunde in der Lage waren, die Leistungen, also den Anteil korrekt gegebener Antworten, der virtuellen Schüler*innen im SKR^{Bio} adäquat einzuschätzen (Rangkomponente: $M_R = 0,39$). Im Vergleich zu Befunden aus der Feldforschung (Südkamp et al. 2012) sowie aus anderen Studien mit dem Simulierten Klassenraum (z. B. Kaiser et al. 2013) fiel die beobachtete Beurteilungsgenauigkeit niedrig aus. Hierbei ist zu bedenken, dass die Lehramtsstudierenden im SKR^{Bio}, anders als etwa bei Kaiser et al. (2013), jede gegebene virtuelle Schüler*innenantwort selbst hinsichtlich der wissenschaftlichen Korrektheit sowie vorliegenden Fehlvorstellungskategorie beurteilen mussten und dementsprechend vergleichsweise stärker kognitiv ausgelastet waren.

6.6.2 Zusammenhang zwischen deklarativem und prozeduralem Wissen

Nach kognitions- und pädagogisch-psychologischen Ansätzen wurde das prozedurale Wissen als Handlungswissen verstanden (z. B. De Jong und Ferguson-Hessler 1996; Lukesch 2001). Auch die naturwissenschaftsdidaktische Forschung schloss sich dieser Konzeptualisierung in zahlreichen Studien an (z. B. Jüttner und Neuhaus 2013; Olszewski et al. 2010, Tepner et al. 2012). Auf dieser Grundlage wurde das gezeigte Wissen der angehenden Biologielehrer*innen im SKR^{Bio} als prozedurales Wissen operationalisiert, da Wissen in einer konkreten Handlungssituation angewendet werden musste, um gegebene virtuelle Schüler*innenantworten sowie die Leistungen am Ende der virtuellen Unterrichtsstunde adäquat einzuschätzen. Die beschriebene Handlungsnähe resultierte ebenso aus der hohen ökologischen Validität, welche der SKR^{Bio} als simulierte Klassenraumumgebung mit sich bringt (Südkamp et al. 2008). Das gezeigte Wissen im Fragebogen wurde als deklaratives Wissen operationalisiert, da zur Beantwortung der Fragen überwiegend auf Wissen zu Fakten, Begriffen und Prinzipien zurückgegriffen werden musste, ohne es in eine konkrete Handlungssituation einzubetten (Krathwohl 2002; Tepner et al. 2012). Die Modellierung verschiedener Wissenstypen sowie die Entwicklung verschiedener Testformate fokussierten bereits einige Forschungsvorhaben (Borowski et al. 2011; Kirschner et al. 2017; Kramer et al. 2021; Krauss et al. 2020; Schmelzing et al. 2013). Auch hier wurde zur Erfassung deklarativen Wissens überwiegend Fragebögen eingesetzt, wohingegen das prozedurale Wissen via videobasierter Erhebungsinstrumente realisiert wurde (Blömeke et al. 2015; Kramer et al. 2021; Krauss et al. 2020). Trotzdem wurde nicht in allen Studien die vorgenommene Differenzierung der Wissenstypen in den Zusammenhangsanalysen explizit berücksichtigt.

Übereinstimmend mit Hypothese 1 zeigten sich überwiegend statistisch signifikante moderate Zusammenhänge zwischen dem deklarativen Wissen im ProwiE-Fragebogen und dem prozeduralen Wissen im SKR^{Bio}. So waren Lehramtsstudierende mit höherem deklarativen Wissen häufiger in der Lage, wissenschaftlich korrekte Antworten zu erkennen, Fehlvorstellungen in Antworten zu identifizieren und dominante Fehlvorstellungen bei den virtuellen Schüler*innen zu diagnostizieren. Basierend auf der evolutionsbiologischen Ausrichtung beider Instrumente sowie den Vorkenntnissen aus der Wissenspsychologie (z. B. Anderson 2001) waren diese Zusammenhänge zu erwarten. Außerdem konnten bereits einige Studien in der bildungswissenschaftlichen und naturwissenschaftsdidaktischen Forschung erste Zusammenhänge zwischen dem deklarativen und prozeduralen Wissen belegen. Beispielsweise zeigte eine Studie von Kramer et al. (2020) einen signifikanten Zusammenhang zwischen dem in einem Fragebogen erfassten FDW und PW angehender Biologielehrkräfte und den gezeigten diagnostischen Aktivitäten in einem videobasierten Messinstrument.

Bei genauerer Betrachtung der einzelnen Fehlvorstellungen zeigte sich, dass höheres deklaratives Wissen augenscheinlich einzig bei lamarckistischen Fehlvorstellungen nicht mit höheren Anteilen korrekter Diagnosen einhergeht. In Anlehnung an theoretische Vorüberlegungen von Shavelson et al. (2005), dass beide Wissenstypen einerseits voneinander abgegrenzt werden können und andererseits sich gegenseitig bedingen, lieferten diese Ergebnisse Hinweise darauf, dass deklaratives und prozedurales biologiespezifisches Wissen zur Evolution bei Lehramtsstudierenden zusammenhängt. Diese Befunde sind besonders dann relevant, wenn das deklarative Wissen angehender Lehrkräfte den Ausgangspunkt bzw. die Voraussetzung für eine erfolgreiche Professionalisierung darstellt (z. B. Anderson 1983; Shavelson et al. 2005). Demnach müsste vorerst deklaratives Wissen im akademischen Diskurs erworben werden, bevor es in konkreten Handlungssituationen angewendet werden kann und, über regelmäßige Übung und Kontextualisierung, prozedurales Wissen entsteht (z. B. Anderson 1982; Blömeke et al. 2010; Schneider und Stern 2010). Ein potenzieller schwacher Zusammenhang mit der Beurteilungsgenauigkeit bei der Leistungsbeurteilung am Ende der Stunde war tendenziell sichtbar, aber bei der gegebenen Stichprobe nicht statistisch signifikant.

Eine differenzierte Betrachtung der Facetten deklarativen Wissens lieferte überraschende Ergebnisse. Im Widerspruch zu den Hypothesen 2 und 3 schienen das deklarative FW eher mit der Diagnose der spezifischen Fehlvorstellungskategorie und das deklarative FDW eher mit Diagnose der wissenschaftlich korrekten Antworten zusammenzuhängen. Zwar waren die Zusammenhangsunterschiede statistisch nicht signifikant, dennoch sind diese Ergebnisse auf Basis bisherigen Forschung schwer zu erklären, da die Diagnose spezifischer Fehlvorstellungen (hier: Wissen über Fehlvorstellungen) eine Kernfacette des FDW ist (z. B. Förtsch et al. 2018; Lee und Luft 2008; Schmelzing et al. 2013) und dementsprechend eher Zusammenhänge mit dem deklarativen FDW zu erwarten gewesen wären. Obwohl beide Instrumente das evolutionsbezogene FW und FDW von Lehramtsstudierenden fokussierten, wird beispielsweise im Fragebogen, anders als im SKR^{Bio}, nicht nur Wissen über Fehlvorstellungen erfasst, sondern auch das Wissen über Instruktionsstrategien (Großschedl et al. 2015b). Hier könnten Gründe für unterschiedlich ausgeprägte Zusammenhänge liegen. Da das deklarative FW größere Zusammenhänge mit dem prozeduralen FDW zu haben scheint, könnte die Annahme gestärkt werden, dass das FW eine wichtige Voraussetzung für das FDW darstellt (z. B. Shulman 1987; Baumert et al. 2010). Durch Replikationsstudien könnten diese ersten Befunde nicht nur theoretisch, sondern auch empirisch fundiert werden.

Kernlimitation der vorliegenden Studie stellt die Konzeptualisierung des deklarativen und prozeduralen Wissens dar. Abhängig von der Forschungsdisziplin sowie

der gewählten Erhebungsinstrumente können die Operationalisierungen verschiedener Wissenstypen stark variieren. Das Wissen über Begriffe, Fakten und Konzepte wird als deklaratives Wissen beschrieben, welches maßgeblich für die Beantwortung der Fragen im ProWiE-Fragebogen verantwortlich ist. Trotzdem wurden beispielsweise Items des ProWiE-Fragebogens durch Großschedl et al. (2015b) der Fähigkeitsdimension „Verstehen & Anwenden“ zugeordnet und somit ein konkreter Anwendungsbezug impliziert, wohingegen In der vorliegenden Studie alle Fragebogenitems dem deklarativen Wissen zugeordnet wurden (vgl. Kramer et al. 2021; Krauss et al. 2020; Schmelzing et al. 2013). Diese Unschärfe bei der Operationalisierung von deklarativem und prozeduralem Wissen könnte die gemessenen Zusammenhangsgrößen beeinflusst haben. Die konkrete Handlungssituation, welche durch den SKR^{Bio} initiiert wird, aktiviert das prozedurale Wissen der angehenden Biologie-Lehramtsstudierenden. Dadurch ist diese simulierte Klassenraumumgebung in der Lage, prozedurales Wissen messen und erfassen zu können. Nichtsdestotrotz zeigen die Terminologien und Konzeptualisierungen der entsprechenden Wissenstypen in anderen bildungswissenschaftlichen und naturwissenschaftsdidaktischen Studien, dass die von uns gewählte Operationalisierung deklarativen und prozeduralen Wissens keine Allgemeingültigkeit besitzt, sondern eine auf das Studiendesign und die Erhebungsinstrumente adaptierte Operationalisierung darstellt. Besonders die Herausforderung, beide Wissenstypen trennscharf voneinander messen und interpretieren zu können, muss in Zukunft mehr Aufmerksamkeit gewidmet werden. Weithin muss genauer untersucht werden, inwiefern der SKR^{Bio} tatsächlich eine konkrete Handlungssituation erzeugt, da dieser Aspekt eine Grundannahme für unsere Konzeptualisierung prozeduralen Wissens ist. Hier könnte die starke Komplexitätsreduktion des SKR^{Bio}, beispielsweise die eindeutige Zuweisung der virtuellen Schüler*innenantworten einer spezifischen Fehlvorstellungskategorie oder die simplifizierende Darstellung des Unterrichtsgesprächs zwischen Lehramtsstudierenden und der virtuellen Schulklasse, die gewünschte ökologische Validität stark limitieren.

6.7 Fazit und Ausblick

Trotz intensiver Forschung im Bereich des Lehrens und Lernens von Evolution (z. B. Gregory 2009) konnten die Ergebnisse der vorliegenden Studie zeigen, dass Lehramtsstudierende der Biologie weiterhin Schwierigkeiten damit haben, spezifische Fehlvorstellungskategorien zur natürlichen Selektion in Schüler*innenantworten zu erkennen. Diese Fokussierung des SKR^{Bio} auf die Diagnose prominenter Fehlvorstellungen limitiert die Facetten eines komplexen Unterrichtsgesprächs, bildet aber die Ausgangslage für das zukünftige unterrichtliche Handeln der Lehrenden. Besonders problematisch werden diese fachdidaktischen Wissenslücken in realen Unterrichtssituationen, wenn es

darum geht, mögliche Anschlusshandlungen zu evaluieren, um das Erlernen und gemeinsame Weiterentwickeln des wissenschaftlich korrekten Evolutionskonzepts zu unterstützen (Ziadie und Andrews 2018). Dementsprechend muss innerhalb der universitären Biologie-Lehramtsausbildung ein starker Fokus auf die Vermittlung der Evolutionstheorie und den damit verbundenen notwendigen fachdidaktischen Kenntnissen gelegt werden, um gegenüber häufig vorkommender Fehlvorstellungen zu sensibilisieren. Hierfür könnte der SKR^{Bio} auch mit entsprechenden Feedbackmaßnahmen weiterentwickelt werden, um über die Funktion als Messinstrument hinaus auch eine Förderung der Diagnosefähigkeit zu bewirken. Dahingehend wurde bereits im SKR^{Bio} ein Feedbackmodul implementiert, welches den Lehramtsstudierenden direkt im Anschluss der Bearbeitung eine summative Rückmeldung zu ihren gezeigten Diagnoseleistungen gibt.

Weiterhin konnte die Studie erste Hinweise liefern, dass das deklarative und prozedurale biologiebezugene Professionswissen zur Evolution angehender Lehrkräfte zusammenhängen. Somit konnte Anschluss an die bisherige kognitionspsychologische Forschung zur Differenzierung und Interaktion deklarativen und prozeduralen Wissens hergestellt werden (z. B. Shavelson et al. 2005). Der Einsatz von simulierten Klassenraumumgebungen scheint eine Möglichkeit zu eröffnen, das im Biologiestudium erworbene deklarative Wissen der Lehramtsstudierenden in einer konkreten Handlungssituation anzuwenden (Badiie und Kaufman, 2015). Um den Lehramtsstudierenden neben den schulpraktischen Phasen weitere Lerngelegenheiten zu ermöglichen, in denen prozedurales Wissen angewendet und trainiert werden kann, bietet die Integration simulierter Klassenraumumgebungen in die universitäre Lehre eine hervorragende Lösung. Hierbei können Lehramtsstudierende orts- und zeitunabhängig ihr Professionswissen auf die Probe stellen und sich auf zukünftige Klassenraumsituationen, z. B. der Diagnose von Schülerantworten zur Evolution im Biologieunterricht, vorbereiten. Besonders die Tatsache, dass prozedurales Wissen eine wichtige Komponente für das professionelle Handeln angehender Lehrkräfte darstellt, sollte zu einer intensiveren Betrachtung sowie Förderung unterschiedlicher Wissenstypen während des naturwissenschaftlichen Lehramtsstudiums führen.

6.8 Literatur

- Abell, S. K. (2007). Research on science teacher knowledge. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (S. 1105–1149). Mahwah, NJ: Lawrence Erlbaum.
- Alters, B. J., & Nelson C. E. (2002). Perspective: Teaching evolution in higher education. *Evolution: International Journal of Organic Evolution*, 56(10), 1891–1901.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge: Harvard University Press.
- Anderson, J. R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum Akademischer Verlag.
- Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, 39(10), 952–978.
- Anderson, J. R., & Lebière, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Andrews, T. M., Kalinowski, S. T., & Leonard, M. J. (2011). "Are humans evolving?" A classroom discussion to change student misconceptions regarding natural selection. *Evolution: Education and Outreach*, 4, 456–466. doi: 10.1007/s12052-011-0343-4
- Baalmann, W., Frerichs, V., Weitzel, H., Gropengießer, H., & Kattmann, U. (2004). Schülervorstellungen zu Prozessen der Anpassung–Ergebnisse einer Interviewstudie im Rahmen der Didaktischen Rekonstruktion. *Zeitschrift für Didaktik der Naturwissenschaften*, 10(1), 7–28.
- Badiee, F., & Kaufman, D. (2015). Design evaluation of a simulation for teacher education. *Sage Open*, 5(2), 1–10.
- Basel, N., Harms, U., Prechtel, H., Weiss, T., & Rothgangel, M. (2014). Students' arguments on the science and religion issue: the example of evolutionary theory and genesis. *Journal of Biological Education*, 48(4), 179–187.
- Baumert, J., & Kunter, M. (2006). Keyword: Professional competence of teachers. *Journal for Educational Science*, 9(4), 469–520.
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.

- Baumert, J. & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (S. 25–48). Springer.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180.
- Beardsley, P. M. (2004). Middle school student learning in evolution: Are current standards achievable? *The American Biology Teacher*, 66(9), 604–613.
- Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, 27(5), 415–427.
- Bizzo, N. M. V. (1994). From down house landlord to Brazilian high school students: what has happened to evolutionary knowledge on the way?. *Journal of Research in Science Teaching*, 31(5), 537–556.
- Blömeke, S., Felbrich, A., Müller, C., Kaiser, G., & Lehmann, R. (2008). Effectiveness of teacher education. *ZDM*, 40(5), 719–734.
- Blömeke, S., Kaiser, G., & Lehmann, R. (Hrsg.). (2010). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. Münster: Waxmann Verlag.
- Blömeke, S., König, J., Suhl, U., Hoth, J., & Döhrmann, M. (2015). Wie situationsbezogen ist die Kompetenz von Lehrkräften? Zur Generalisierbarkeit der Ergebnisse von videobasierten Performanztests. *Zeitschrift für Pädagogik*, 61(3), 310–327.
- Borowski, A., Kirschner, S., Liedtke, S. & Fischer, H. E. (2011). Vergleich des Fachwissens von Studierenden, Referendaren und Lehrenden in der Physik. *PhyDid- Physik und Didaktik in Schule und Hochschule.*, 10(1), 1–9.
- Borowski, A., Neuhaus, B. J., Tepner, O., Wirth, J., Fischer, H. E., Leutner, D., Sandmann, A., & Sumfleth, E. (2010). Professionswissen von Lehrkräften in den Naturwissenschaften (ProwiN)–Kurzdarstellung des BMBF-Projekts. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 341–349.
- BouJaoude, S., Asghar, A., Wiles, J. R., Jaber, L., Sameddine, D., & Alters, B. (2011). Biology professors' and teachers' positions regarding biological evolution and evolution education in a Middle Eastern society. *International Journal of Science Education*, 33(7), 979–1000.

- Bromme, R. (2001). Teacher Expertise. In N. J. Smelser, P. B. Baltes, & F. E. Weinert (Hrsg.), *International Encyclopedia of the Behavioral Sciences: Education* (S. 15459–15465). London: Pergamon.
- Bromme, R. (2008). Lehrerexpertise. In W. Schneider, & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 159–167). Göttingen: Hogrefe.
- Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education*, 68(4), 493–503.
- Brunner, M., Anders, Y., Hachfeld, A., & Krauss, S. (2011). 10 Diagnostische Fähigkeiten von Mathematiklehrkräften. *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*, 215–233.
- Byrne, J., Grace, M., & Hanley, P. (2009). Children's anthropomorphic and anthropocentric ideas about micro-organisms: Educational research. *Journal of biological education*, 44(1), 37–43.
- De Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105–113.
- Demastes, S. S., Good, R. G., & Peebles, P. (1996). Patterns of conceptual change in evolution. *Journal of Research in Science Teaching*, 33(4), 407–431.
- Demastes, S. S., Settlage Jr, J., & Good, R. (1995). Students' conceptions of natural selection and its role in evolution: Cases of replication and comparison. *Journal of Research in Science Teaching*, 32(5), 535–550.
- van Driel, J. H., Verloop, N., & De Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, 35(6), 673–695.
- Edelmann, W. (2000). *Lernpsychologie*. Kempten: Kösel-Verlag.
- Endler, J. A. (1986). *Natural selection in the wild*. Princeton, NJ: Princeton University Press.
- Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. *Review of Research in Education*, 20(1), 3–56. doi:10.3102/0091732X020001003
- Ferrari, M., & Chi, M. T. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education*, 20(10), 1231–1256.
- Fiedler, K., Freytag, P. & Unkelbach, C. (2007). Pseudo-contingencies in a simulated classroom. *Journal of Personality and Social Psychology*, 92, 665–677.

- Fiedler, K., Walther, E., Freytag, P. & Plessner, H. (2002). Judgment biases in a simulated classroom – A cognitive-environmental approach. *Organizational Behavior and Human Decision Processes*, 88, 527–561.
- Fischer, H. E., Borowski, A., & Tepner, O. (2012). Professional knowledge of science teachers. In *Second International Handbook of Science Education* (S. 435–448). Dordrecht: Springer.
- Friedrichsen, P. J., Abell, S. K., Pareja, E. M., Brown, P. L., Lakford, D. M., & Volkmann, M. J. (2009). Does teaching experience matter? Examining biology teachers' prior knowledge for teaching in an alternative certification program. *Journal of Research in Science Teaching*, 46(4), 357–383.
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, 49(9), 1181–1210.
- Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M., Girwidz, R., Obersteiner, A., ... & Seidel, T. (2018). Systematizing professional knowledge of medical doctors and teachers: development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences*, 8(4), 207–225.
- Graf, D. & Hamdorf, E. (2011). Evolution: Verbreitete Fehlvorstellungen zu einem zentralen Thema. In: Dreesmann, D., Graf, D. & Witte, K. (Hrsg.): *Evolutionsbiologie: Moderne Themen für den Unterricht*, (S. 25–41). Heidelberg: Springer Akademischer Verlag.
- Gregory, T. R. (2009). Understanding Natural selection. Essential concepts and common misconceptions. *Evolution: Education and Outreach*, 2(2), 156–175.
- Gresch, H., & Martens, M. (2019). Teleology as a tacit dimension of teaching and learning evolution: A sociological approach to classroom interaction in science education. *Journal of Research in Science Teaching*, 56(3), 243–269.
- Großschedl, J., Harms, U., Kleickmann, T., & Glowinski, I. (2015a). Preservice biology teachers' professional knowledge: Structure and learning opportunities. *Journal of Science Teacher Education*, 26(3), 291–318.
- Großschedl, J., Konnemann, C., & Basel, N. (2014). Pre-service biology teachers' acceptance of evolutionary theory and their preference for its teaching. *Evolution: Education and Outreach*, 7(18), 1–16.
- Großschedl, J., Neubrand, C., Kirchner, A., Oppermann, L., Basel, N., & Gantner, S. (2015b). Entwicklung und Validierung eines Testinstruments zur Erfassung des

- evolutionsbezogenen Professionswissens von Lehramtsstudierenden (ProwiE). *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 173–185.
- Großschedl, J., Seredszus, F., & Harms, U. (2018). Angehende Biologielehrkräfte: evolutionsbezogenes Wissen und Akzeptanz der Evolutionstheorie. *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 51–70.
- Großschedl, J., Welter, V., & Harms, U. (2019). A new instrument for measuring pre-service biology teachers' pedagogical content knowledge: The PCK-IBI. *Journal of Research in Science Teaching*, 56(4), 402–439.
- Gruber, H., Mandl, H., & Renkl, A. (2000). Was lernen wir in Schule und Hochschule: Träges Wissen?. In H. Mandl & J. Gerstenmaier K. (Hrsg.), *Die Kluft zwischen Wissen und Handeln: Empirische und theoretische Lösungsansätze* (S. 139–156). Göttingen: Hogrefe.
- Gruber, H., & Renkl, A. (2000). Die Kluft zwischen Wissen und Handeln: Das Problem des trägen Wissens. In G. H. Neuweg (Hrsg.), *Wissen – Können – Reflexion. Ausgewählte Verhältnisbestimmungen* (S. 155–174). Innsbruck: Studienverlag.
- Hammann M, Nehm R. (2020). Teleology and evolution education: introduction to the special issue. *Evolution: Education and Outreach*, 13(16), 1–5.
- Heitzman, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M., ... & Fischer, F. (2019). Facilitating Diagnostic Competences in Simulations: A Conceptual Framework and a Research Agenda for Medical and Teacher Education. *Frontline Learning Research*, 7(4), 1–24.
- Helmke, A., Hosenfeld, I. & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften. In R. Arnold & C. Giese (Hrsg.), *Schulleitung und Schulentwicklung* (S. 119–144). Hohengehren: Schneider-Verlag.
- Jensen, M. S., & Finley, F. N. (1996). Changes in students' understanding of evolution resulting from different curricular and instructional strategies. *Journal of Research in Science Teaching*, 33(8), 879–900.
- Jiménez-Aleixandre, M. P. (1992). Thinking about theories or thinking with theories? A classroom study with natural selection. *International Journal of Science Education*, 14(1), 51–61.
- Jungwirth, E. (1977). Should natural phenomena be described teleologically or anthropomorphically?—a science educator's view. *Journal of Biological Education*, 11(3), 191–196.

- Jüttner, M., Boone, W., Park, S., & Neuhaus, B. J. (2013). Development and use of a test instrument to measure biology teachers' content knowledge (CK) and pedagogical content knowledge (PCK). *Educational Assessment, Evaluation and Accountability*, 25(1), 45–67.
- Jüttner, M., & Neuhaus, B. J. (2013). Das Professionswissen von Biologielehrkräften—Ein Vergleich zwischen Biologielehrkräften, Biologen und Pädagogen. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 31–49.
- Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction*, 28, 73–84.
- Kaiser, J., Südkamp, A., & Möller, J. (2016). The effects of student characteristics on teachers' judgment accuracy: Disentangling ethnicity, minority status, and achievement. *Journal of Educational Psychology*, 109, 881–888. doi:10.1037/edu0000156
- Kalinowski, S. T., Leonard, M. J., & Taper, M. L. (2016). Development and validation of the conceptual assessment of natural selection (CANS). *CBE—Life Sciences Education*, 15(4), 1–11.
- Kallery, M., & Psillos, D. (2004). Anthropomorphism and animism in early years science: Why teachers use them, how they conceptualise them and what are their views on their use. *Research in Science Education*, 34(3), 291–311.
- Kampourakis, K., & Zogza, V. (2008). Students' intuitive explanations of the causes of homologies and adaptations. *Science & Education*, 17(1), 27–47.
- Kirschner, S., Sczudlek, M., Tepner, O., Borowski, A., Fischer, H. E., Lenske, G., ... & Wirth, J. (2017). Professionswissen in den Naturwissenschaften (ProwiN). In *Entwicklung von Professionalität pädagogischen personals* (S. 113–130). Wiesbaden: Springer.
- Kleickmann, T., Grosschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., Koller, O., Kroger, J., Lindmeier, A., Loch, C., Mahler, D., Moller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V., & Zimmermann, F. (2014). Professionswissen angehender Lehrkräfte mit mathematisch-naturwissenschaftlichen Fachern – Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft*, 42(3), 280–288.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., ... Vollmer, H. J. (2007). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. In BMBF (Hrsg.), *Bildungsforschung* (Bd. 1). Bonn, Berlin: BMBF.

- Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic competence of teachers: A process model that accounts for diagnosing learning behavior tested by means of a case scenario. *Teaching and Teacher education*, 30, 38-46.
- KMK [Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland]. (2005). Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss—Beschluss vom 16.12.2004. München: Luchterhand.
- Kramer, M., Förtsch, C., Boone, W. J., Seidel, T., & Neuhaus, B. J. (2021). Investigating Pre-Service Biology Teachers' Diagnostic Competences: Relationships between Professional Knowledge, Diagnostic Activities, and Diagnostic Accuracy. *Education Sciences*, 11(3), 89.
- Kramer, M., Förtsch, C., Stürmer, J., Förtsch, S., Seidel, T. & Neuhaus, B. J. (2020). Measuring biology teachers' professional vision: Development and validation of a video-based assessment tool. *Cogent Education*, 7(1).
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212–218.
- Krauss, S., Bruckmaier, G., Lindl, A., Hilbert, S., Binder, K., Steib, N., & Blum, W. (2020). Competence as a continuum in the COACTIV study: the “cascade model”. *ZDM*, 52(2), 311-327.
- König, J. (2010). Lehrerprofessionalität - Konzepte und Ergebnisse der internationalen und deutschen Forschung am Beispiel fachübergreifender, pädagogischer Kompetenzen. In J. König, & B. Hofmann (Hrsg.), *Professionalität von Lehrkräften – Was sollen Lehrkräfte im Lese- und Schreibunterricht wissen und können?*. (S. 40–105). Berlin: DGLS.
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527.
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A., & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teaching and Teacher Education*, 38, 76–88.
- Lee, E., & Luft, J. A. (2008). Experienced secondary science teachers' representation of pedagogical content knowledge. *International Journal of Science Education*, 30(10), 1343–1363.
- Lucero, M. M., Delgado, C., & Green, K. (2019). Elucidating High School Biology Teachers' Knowledge of Students' Conceptions Regarding Natural Selection. *International Journal of Science and Mathematics Education*, 1–21.

- Lukesch, H. (2001). *Psychologie des Lernens und Lehrens*. Regensburg: Roderer.
- Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In *Examining pedagogical content knowledge* (S. 95–132). Dordrecht: Springer.
- Mahler, D., Großschedl, J., & Harms, U. (2017). Using doubly latent multilevel analysis to elucidate relationships between science teachers' professional knowledge and students' performance. *International Journal of Science Education*, 39(2), 213–237.
- Mayr, E. (1982). *The growth of biological thought: Diversity, evolution and inheritance*. Cambridge, MA: Harvard University Press.
- McVaugh, N. K., Birchfield, J., Lucero, M. M., & Petrosino, A. J. (2011). Evolution education: Seeing the forest for the trees and focusing our efforts on the teaching of evolution. *Evolution: Education and Outreach*, 4(2), 286–292.
- Moore, R., Mitchell, G., Bally, R., Inglis, M., Day, J., & Jacobs, D. (2002). Undergraduates' understanding of evolution: ascriptions of agency as a problem for student learning. *Journal of Biological Education*, 36(2), 65–71.
- Nationale Akademie der Wissenschaften Leopoldina. (2017). Evolutionsbiologische Bildung in Schule und Universität. Deutsche Akademie der Naturforscher Leopoldina e.V. Nationale Akademie der Wissenschaften, Halle (Saale).
- National Research Council [NRC]. (2012). Thinking evolutionarily: Evolution education across the life sciences: Summary of a convocation. Washington, DC: The National Academies Press.
- Nehm, R. H., Beggrow, E. P., Opfer, J. E., & Ha, M. (2012). Reasoning about natural selection: diagnosing contextual competency using the ACORNS instrument. *The American Biology Teacher*, 74(2), 92–98.
- Nehm, R. H., Poole, T. M., Lyford, M. E., Hoskins, S. G., Carruth, L., Ewers, B. E., & Colberg, P. J. (2009). Does the segregation of evolution in biology textbooks and introductory courses reinforce students' faulty mental models of biology and evolution?. *Evolution: Education and Outreach*, 2(3), 527–532.
- Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, 57(3), 263–272.
- Nehm, R. H., & Schonfeld, I. S. (2007). Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools?. *Journal of Science Teacher Education*, 18(5), 699–723.

- Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, 45(10), 1131–1160.
- NGSS Lead States (2013). Next generation science standards: for states, by states. Washington, DC: The National Academies Press.
- Olander, C. (2012). Teaching biological evolution-internal and external evaluation of learning outcomes. *Nordic Studies in Science Education*, 5(2), 171–184.
- Olszewski, J., Neumann, K., & Fischer, H. E. (2010). Measuring physics teachers' declarative and procedural PCK. *Contemporary science education research: Teaching*. Ankara: PegemAkademi, 87–94.
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: knowing what students know about evolution. *Journal of Research in Science Teaching*, 49(6), 744–777.
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261–284.
- Perry, M. (1991). Learning and transfer: instructional conditions and conceptual change. *Cognitive Development*, 6, 449–468.
- Prinou, L., Halkia, L., & Skordoulis, C. (2008). What conceptions do Greek school students form about biological evolution?. *Evolution: Education and Outreach*, 1(3), 312–317.
- Riese, J., Gramzow, Y., & Reinhold, P. (2017). Die Messung fachdidaktischen Wissens bei Anfängern und Fortgeschrittenen im Lehramtsstudiengang Physik. *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 99–112.
- Riese, J. & Reinhold, P. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 167–187.
- Rittle-Johnson, B., Siegler, R. S. & Wagner Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: an interative process. *Journal of Educational Psychology*, 93, 346–362.
- Rittle-Johnson, B. & Wagner Alibali, M. (1999). Conceptual and procedural knowledge of mathematics: does one lead to the other?. *Journal of Educational Psychology*, 91, 175–189.

- Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal*, 50(5), 1020–1049.
- Santagata, R., & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM*, 43(1), 133–145. <https://doi.org/10.1007/s11858-010-0292-3>
- Schmelzing, S., van Driel, J. H., Jüttner, M., Brandenbusch, S., Sandmann, A., & Neuhaus, B. J. (2013). Development, evaluation, and validation of a paper-and-pencil test for measuring two components of biology teachers' pedagogical content knowledge concerning the "cardiovascular system". *International Journal of Science and Mathematics Education*, 11, 1369–1390.
- Schmelzing, S., Wüsten, S., Sandmann, A., & Neuhaus, B. (2010a). Fachdidaktisches Wissen und Reflektieren im Querschnitt der Biologielehrerbildung. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 189–207.
- Schmelzing, S., Wüsten, S., Sandmann, A., & Neuhaus, B. (2010b). Measuring declarative and reflective components of biology teachers' pedagogical content knowledge. *Contemporary science education research: pre service and inservice teacher education*. Ankara, Turkey: Pegem Akademi, 71–77.
- Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology*, 46(1), 178–192.
- Schrader, F.-W. (2006). Diagnostische Kompetenz von Eltern und Lehrern. In D.H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 95–100). Weinheim: Beltz.
- Schrader, F.-W. & Helmke, A. (2001). Alltägliche Leistungsbeurteilung durch Lehrer. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 45–58). Weinheim: Beltz.
- Settlage Jr, J. (1994). Conceptions of natural selection: A snapshot of the sense-making process. *Journal of Research in Science Teaching*, 31(5), 449–457.
- Shavelson, R. J., Ruiz-Primo, M. A. & Wiley, E. W. (2005). Windows into the mind. *Higher Education*, 49(4), 413–430. doi: 10.1007/s10734-004-9448-9
- Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, 124, 209–215.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review*, 57(1), 1–23.

- Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education and Outreach*, 1(2), 189–195.
- Smith, C. H., & Beccaloni, G. (Hrsg.) (2008). *Natural Selection and Beyond. The intellectual legacy of Alfred Russel Wallace*. New York: Oxford University Press.
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19, 85–95
- Steiner, G. (2001). Lernen als Wissenserwerb. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 163–202). Weinheim: PVU.
- Stover, S. K., & Mabry, M. L. (2007). Influences of Teleological and Lamarckian Thinking on Student Understanding of Natural Selection. *Bioscience*, 33(1), 11–18.
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762.
- Südkamp, A., Möller, J., & Pohlmann, B. (2008). Der Simulierte Klassenraum: Eine experimentelle Untersuchung zur diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 22(34), 261–276.
- Tamir, P., & Zohar, A. (1991). Anthropomorphism and teleology in reasoning about biological phenomena. *Science Education*, 75(1), 57–67.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S.,... Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- Tibell, L. A., & Harms, U. (2017). Biological principles and threshold concepts for understanding natural selection. *Science and Education*, 26(7–9), 953–973.
- Voss, T. & Kunter, M. (2011). Pädagogischpsychologisches Wissen von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 193–214). Münster: Waxmann.
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical and psychological knowledge: test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969. <https://doi.org/10.1037/a0025125>.

- Wollenschläger, M., Hattie, J., Machts, N., Möller, J., & Harms, U. (2016). What makes rubrics effective in teacher-feedback? Transparency of learning goals is not enough. *Contemporary Educational Psychology*, 44, 1–11.
- Yip, D. Y. (1998). Identification of misconceptions in novice biology teachers and remedial strategies for improving biology learning. *International Journal of Science Education*, 20(4), 461–477.
- Zabel, J., & Gropengiesser, H. (2011). Learning progress in evolution theory: climbing a ladder or roaming a landscape?. *Journal of biological education*, 45(3), 143–149.
- Ziadie, M. A., & Andrews, T. C. (2018). Moving evolution education forward: a systematic analysis of literature to identify gaps in collective knowledge for teaching. *CBE—Life Sciences Education*, 17(1), 1–10.

7 ZUSAMMENFASSUNG DER ERGEBNISSE DER DREI STUDIEN

7.1 Studie 1: Measuring Biology Trainee Teachers' Professional Knowledge about Evolution – Introducing the Student Inventory

Das Ziel der ersten Studie war es, die biologiespezifischen und generischen Facetten des prozeduralen professionellen Wissens von LiV des Faches Biologie zu untersuchen, und erste Hinweise auf den Zusammenhang zwischen deklarativem und prozeduralem Wissen zu gewinnen. Inhaltlich ging es um das Wissen zur Evolution, speziell um den Prozess der natürlichen Selektion. Um Einblicke in die diagnostischen Aktivitäten der LiV auf Basis ihres Professionswissens zu erhalten, wurde das digitale Instrument – das Schülerinventar (SI) – weiterentwickelt und mit virtuellen Schülerklausuren zur Evolution gefüllt, welche sich in der Qualität (wissenschaftlich vs. fehlvorstellungsbeinhaltende Schülertexte) unterschieden. Als Grundlage dienten Studien zu spezifischen Fehlvorstellungskategorien zum Prozess der natürlichen Selektion. Innerhalb der Schülerklausuren wurde zusätzlich zu den Schülertexten vorab jeweils eine weitere leistungsirrelevante Information (Ergebnisse aus einem Multiple-Choice Test zur Evolution) gegeben, um zu beobachten, ob diese Vorinformation die Beurteilung der Schülertexte systematisch verzerrt. Das SI wurde von $N = 27$ LiV bearbeitet, wobei jeweils insgesamt sechs virtuelle Schülerklausuren beurteilt werden mussten. Insgesamt sind 162 beurteilte virtuelle Schülerklausuren in die Analyse eingeflossen.

Die Ergebnisse zeigten, dass die Lehrerinnen und Lehrer im Vorbereitungsdienst in der Lage waren, zwischen wissenschaftlichen und nicht-wissenschaftlichen Denkweisen in den virtuellen Schülertexten zu unterscheiden. Problematisch war hingegen die Diagnose von spezifischen Fehlvorstellungskategorien (anthropomorph, teleologisch) innerhalb der virtuellen Schülertexte. Hierbei wurden die anthropomorphen Fehlvorstellungen deutlich häufiger erkannt als geäußerte teleologische Fehlvorstellungen. Außerdem konnte gezeigt werden, dass leistungsirrelevante Informationen einen Einfluss auf die Beurteilungsgenauigkeit der LiV hatten. Das heißt, virtuelle Schülertexte gleicher Qualität wurden besser bewertet, wenn vorab ein gutes Multiple-Choice Testergebnis präsentiert wurde und vice versa. Somit konnte gezeigt werden, dass die Urteile der LiV systematisch verzerrt wurden. Durch die Integration eines kurzen Wissenstests zur Evolution mit zwölf Fragen in das SI zur Erfassung des deklarativen Wissens der LiV konnten erste Hinweise darauf geben werden, dass das deklarative und prozedurale biologiebezugene Professionswissen (FW und FDW) zusammenhängen.

7.2 Studie 2: Simulated Classroom Biology (SCR^{Bio}) – a Simulated Classroom Environment for Assessing the Action-Oriented Professional Knowledge of Pre-Service Teachers About Evolution

Die zweite Studie diente der Erprobung und Validierung des neu entwickelten SKR^{Bio}. Hier stand die Frage im Mittelpunkt, inwiefern diese simulierte Klassenraumumgebung eine valide Testwertinterpretation des prozeduralen FDW zulässt. Zur theoretischen Konzeptualisierung des FDW wurde daher das aus der fachdidaktischen Forschung stammende *RCM (Refined Consensus Model of PCK)* fokussiert, welches in deklarative und prozedurale Facetten des FDW differenziert. Anfänglich wurde der SKR^{Bio} mit evolutionsbiologischen Fragestellungen zum Prozess der natürlichen Selektion verschiedener Organismengruppen ($N = 27$ Fragen) sowie entsprechender virtueller Schülerantworten ($N = 476$) erweitert. Die Antworten der virtuellen Schülerinnen und Schüler artikulierten entweder eine wissenschaftlich korrekte Denkweise oder eine spezifische Fehlvorstellungskategorie (anthropomorph, lamarckistisch, teleologisch). Ein durchgeführtes Expertenrating ($N = 10$) prüfte im ersten Schritt die Evidenz basierend auf dem Testinhalt. Um Evidenz basierend auf Gruppenvergleichen zu generieren, wurde die Stichprobe ($N = 76$ Biologie-Lehramtsstudierende) zufällig in zwei Experimentalgruppen geteilt, welche entweder einen fachdidaktisch-bezogenen Prompt oder einen alternativen Prompt erhielten. Es wurde erwartet, dass die Expertinnen und Experten die wissenschaftlich korrekten Denkweisen sowie die spezifischen Fehlvorstellungskategorien innerhalb der entwickelten virtuellen Schülerantworten korrekt zuweisen können, und, dass der fachdidaktisch-bezogene Prompt die Diagnose spezifischer Fehlvorstellungskategorien der Lehramtsstudierenden unterstützt (FDW), aber keinen Einfluss auf die Beurteilungsgenauigkeit hat (PW).

Die Interrater-Übereinstimmung des Expertenratings zehn unabhängiger Expertinnen und Experten bestätigten die konzeptionelle Entwicklung der Fragen sowie die Erstellung der virtuellen Schülerantworten, welche eindeutig eine intendierte spezifische Fehlvorstellungskategorie oder wissenschaftliche Denkweise artikulierten. Weiterhin konnte gezeigt werden, dass der SKR^{Bio} sensitiv auf das prozedurale FDW der Lehramtsstudierenden reagierte, da die Experimentalgruppe, welche den fachdidaktisch-bezogenen Prompt bekamen, eine höhere Diagnoseleistung der spezifischen artikulierten Fehlvorstellungskategorie in den virtuellen Schülerantworten zeigten als die Kontrollgruppe. Im Gegensatz dazu hatte weder der fachdidaktisch-bezogene noch der alternative Prompt einen Einfluss auf die Beurteilungsgenauigkeit (PW) der Lehramtsstudierenden. Das unterstützt die Annahme, dass das Wissen, welches zur Diagnose spezifischer Fehlvorstellungskategorien (FDW) ein anderes ist als das Wissen, welches für die allgemeine pädagogische Beurteilungsgenauigkeit notwendig ist (PW). Die Betrachtung beider Evidenzkriterien (Evidenz basierend auf dem Testinhalt und Evidenz

basierend auf Gruppenvergleichen) bestätigten die valide Testwertinterpretation des prozeduralen FDW (*ePCK*) bezogen auf das Wissen über Schülerfehlvorstellungen innerhalb des SKR^{Bio}.

7.3 Studie 3: Der Simulierte Klassenraum – Erfassung deklarativen und prozeduralen Wissens bei Lehramtsstudierenden der Biologie

Nach Validierung des SKR^{Bio} lag der Fokus in Studie 3 auf der Untersuchung der Zusammenhänge zwischen deklarativen und prozeduralen Facetten des Professionswissens zur Evolution bei Lehramtsstudierenden der Biologie ($N = 51$). Basierend auf theoretischen Vorkenntnissen sowie empirischen Befunden aus der Wissenspsychologie ist bekannt, dass das deklarative sowie prozedurale Wissen miteinander interagieren. Zur Erfassung beider Wissenstypen wurde innerhalb der Studie eine Kombination zweier Instrumente – ein Fragebogen und der SKR^{Bio} – eingesetzt. Das gezeigte FW und FDW im Fragebogen wurde als deklaratives Wissen operationalisiert, da die Lehramtsstudierenden zur Beantwortung auf ihr Wissen zu Fakten, Begriffen und Prinzipien zurückgreifen mussten. Als prozedurales Wissen wurde das gezeigte FW, FDW und PW im SKR^{Bio} operationalisiert, da die entsprechenden Wissensbereiche hier in eine konkrete Handlungssituation eingebettet waren. Bezüglich der Zusammenhänge des deklarativen und prozeduralen Wissens bei Lehramtsstudierenden wurde angenommen, dass sich stärkere Zusammenhänge innerhalb eines Wissensbereichs zeigen als zwischen den Wissensbereichen.

Die Ergebnisse ergaben überwiegend schwache bis moderate Zusammenhänge zwischen dem gezeigten deklarativen Wissen im Fragebogen und dem gezeigten prozeduralen Wissen im SKR^{Bio} der Lehramtsstudierenden. Hierbei waren Zusammenhänge zwischen dem deklarativen Wissen und der Diagnose wissenschaftlicher korrekter virtueller Schülerantworten, spezifischer Fehlvorstellungen in virtuellen Schülerantworten und der am häufigsten verwendeten Fehlvorstellungskategorie der virtuellen Schülerinnen und Schüler im SKR^{Bio} zu beobachten. Entgegen unserer Annahmen fanden sich keine signifikanten Zusammenhänge innerhalb des deklarativen und prozeduralen FW sowie innerhalb des deklarativen und prozeduralen FDW. Weiterhin wurde in der Studie festgestellt, dass die Zusammenhänge zwischen dem deklarativen Wissen und dem prozeduralen PW nur knapp nicht signifikant waren. Zusammenfassend liefern diese Ergebnisse erste Hinweise darauf, dass das deklarative und prozedurale biologiebezogene Professionswissen zur Evolution von Lehramtsstudierenden zusammenhängen.

8 DISKUSSION UND AUSBLICK

8.1 Übergreifende Diskussion

Die drei im Rahmen dieser Dissertation durchgeführten Studien tragen dazu bei, die Forschungslücke im Bereich der Verfügbarkeit adäquater simulierter Klassenraumumgebungen zur Erfassung von Facetten prozeduralen Professionswissens zur Evolution sowie deren Interaktion mit Facetten deklarativen professionellen Wissens ein Stück weit zu schließen. Nach der umfangreichen Entwicklung und Validierung des SKR^{Bio} zum evolutionsbiologischen Prozess der natürlichen Selektion und erster Erkenntnisse zu Facetten des prozeduralen FW, FDW und PW innerhalb dieser simulierten Klassenraumumgebung wurden anschließend Zusammenhänge zwischen dem deklarativen und dem prozeduralen Wissen untersucht. Die Ergebnisse der einzelnen Studien werden ausführlich in den Manuskripten diskutiert (Kapitel 4 bis 6). In diesem Kapitel werden die Studien nun übergreifend diskutiert und im Zusammenhang betrachtet. Hierbei wird sich auf drei Schwerpunkte konzentriert: Erstens wird die Entwicklung der simulierten Klassenraumumgebung – der SKR^{Bio} – fokussiert (Kapitel 8.1.1). Zweitens wird die Operationalisierung und Differenzierung des deklarativen und prozeduralen Wissens in den Studien diskutiert, um die Befunde in die bisherige kognitionspsychologische sowie naturwissenschaftliche Forschung einzuordnen (Kapitel 8.2.2). Drittens wird die Erfassung des prozeduralen FDW innerhalb der Facette Wissen über Schülerfehlvorstellungen mit dem SKR^{Bio} diskutiert (Kapitel 8.2.3).

8.1.1 Entwicklung des SKR^{Bio}

Zur Durchführung der Studie 2 und 3 wurde der Simulierte Klassenraum, welcher bereits in psychologischen Studien Gegenstand war, für die Beantwortung biologiedidaktischer Fragestellungen weiterentwickelt. Den theoretischen Rahmen für diese inhaltliche Weiterentwicklung lieferten die Forschungen zum Lehren und Lernen der Evolution. Anvisiert wurde die Integration von evolutionsbezogenen Fragestellungen zur natürlichen Selektion unterschiedlicher Organismen sowie dazugehörigen virtuellen Schülerantworten in Textform. In den vergangenen Jahrzehnten identifizierte die fachdidaktische Forschung zahlreiche Fehlvorstellungen von Schülerinnen und Schülern, welche als Erklärungsgerüst evolutionsbiologischer Mechanismen und Prozesse genutzt werden (z. B. Graf und Hamdorf, 2011). Hierbei kristallisierten sich vor allem drei Fehlvorstellungskategorien heraus, welche besonders häufig durch Schülerinnen und Schüler geäußert werden – die anthropomorphe, die teleologische und die lamarckistische Fehlvorstellung (z. B. Bishop und Anderson, 1990; Nehm et al., 2009). Dementsprechend waren diese spezifischen

Fehlvorstellungskategorien der konzeptionelle Rahmen für die Entwicklung der virtuellen Schülerantworten und wurden folglich in allen drei Studien zur Erfassung des FDW genutzt. Die fachdidaktische Forschung konnte zudem zeigen, dass sich das Wissen zur natürlichen Selektion je nach adressierter Organismengruppe unterscheidet (z. B. Opfer et al., 2012; Großschedl et al., 2018). Entsprechend wurden bei der Fragenentwicklung für den SKR^{Bio} verschiedene Organismengruppen berücksichtigt, um diese Disparitäten des FDW besser differenzieren zu können. Innerhalb der Fragestellungen wurde weiter zwischen Merkmalsverlust oder –erwerb unterschieden, da es auch hier Erkenntnisse gibt, dass die Merkmalsausprägung einen Einfluss auf das Verständnis der Evolution haben kann (Nehm et al., 2012).

Da valide Instrumente die Voraussetzung für eine evidenzbasierte Testwertinterpretation darstellen, wurde der SKR^{Bio} in Studie 2 hinsichtlich verschiedener Evidenzkriterien überprüft. Sowohl ein Expertenrating (d.h. Evidenz basierend auf dem Testinhalt) als auch eine experimentelle Variation zweier Gruppen (Evidenz basierend auf Gruppenvergleichen) bestätigten die Hypothesen, dass der SKR^{Bio} geeignet ist, prozedurales FDW adäquat zu erfassen. Hierbei ist zu beachten, dass dem SKR^{Bio} nicht pauschal Validität unterstellt werden darf, sondern nur hinsichtlich der Interpretation von Testergebnissen zum Wissen über Schülerfehlvorstellungen (FDW) angehender Lehrkräfte (Kane, 2013). Jedoch bestätigte die durchgeführte Validierungsstudie die theoretisch und empirisch fundierte Entwicklung des SKR^{Bio} und legt den Grundstein für potenzielle Folgestudien mit weiteren fachdidaktischen Fragestellungen. Darüber hinaus trägt die valide Testwertinterpretation im SKR^{Bio} dazu bei, die Lücke zwischen den verfügbaren simulierten Klassenraumumgebungen zu schließen. Diese adressierten bisher nur prozedurales PW der angehenden Lehrerinnen und Lehrer. Nun ist auch die Messung des prozeduralen biologiebezogenen FDW zur Evolution möglich.

Viele der Vorteile, die bereits in der Forschung zu simulierten Klassenraumumgebungen beschrieben wurden, sind auch für den SKR^{Bio} gültig. Beispielsweise offeriert der SKR^{Bio} eine komplexitätsreduzierte Lernumgebung mit kontrollierbaren Variablen, in welcher die kognitive Belastung reguliert werden kann (z. B. Gurvitch und Metzler, 2009). Weiterhin können falsche Handlungen im SKR^{Bio} keinen negativen Einfluss auf reale Schülerinnen und Schüler haben, wodurch Ängste vor Konsequenzen oder dem Scheitern genommen werden können (z. B. Hixon und So, 2009). Besonders hervorzuheben ist die Handlungsnähe im SKR^{Bio}, indem die Nutzerinnen und Nutzer in die Rolle der Lehrkraft schlüpfen und eine virtuelle Unterrichtsstunde durchführen können (Veletsianos et al., 2010). Erst diese Eingebundenheit der Nutzerinnen und Nutzer in das virtuelle Geschehen und die Notwendigkeit, Wissen in eine konkrete Handlungssituation einzubringen, ermöglicht die Erfassung prozeduralen professionellen Wissens.

8.1.2 Operationalisierung und Differenzierung deklarativen und prozeduralen Wissens

Neuere Konzeptualisierungen des professionellen Wissens von Lehrerinnen und Lehrern adressieren immer häufiger die Anwendung des Professionswissens in konkreten unterrichtlichen Handlungssituationen (z. B. Baumert und Kunter, 2011; Carlson und Daehler, 2019). Folglich werden Instrumente benötigt, die das in den Modellen beschriebene handlungsnahe bzw. prozedurale Wissen erfassen und messen können. Hierfür muss vorerst theoretisch geklärt werden, inwiefern verschiedene Wissenstypen differenziert und operationalisiert werden können. Vor dem Hintergrund der Kognitionsforschung besteht ein Konsens über die Differenzierung in deklaratives und prozedurales Wissen, welchem sich in der vorliegenden Dissertation angeschlossen wurde. Hierbei wird das deklarative Wissen als Wissen über Fakten, Begriffe und Prinzipien verstanden (z. B. Krathwohl, 2002). Sobald Wissen in einer spezifischen Handlung angewendet wird, ist das prozedurale Wissen relevant (z. B. Tepner et al., 2012). Trotzdem bleibt häufig unklar, wie sich diese Wissenstypen in unterschiedlichen Situationen operationalisieren lassen und welche Zusammenhänge zwischen ihnen bestehen. Dementsprechend divergieren die Konzeptualisierungen in der kognitions- und bildungspsychologischen sowie naturwissenschaftlichen Forschung.

In den durchgeführten Studien wurde sich der Konzeptualisierung angeschlossen, wonach sich das prozedurale Wissen direkt in Handlungen operationalisiert und dementsprechend auch als Handlungswissen verstanden werden kann (z. B. De Jong und Ferguson-Hessler, 1996). Folglich ist die zentrale Voraussetzung zur Erfassung prozeduralen Wissens, dass die Erhebungsinstrumente eine konkrete Handlungssituation suggerieren können und gleichzeitig eine gewisse Unmittelbarkeit der durchzuführenden Handlung erzeugen. In Studie 1 wurde mit der Integration von Schülerklausuren in das SI versucht, eine konkrete Handlungssituation zu generieren, in welcher die LiV verschiedene virtuelle Schülerleistungen beurteilen mussten. Hierbei knüpft die Tätigkeit der Korrektur von Schülerklausuren unmittelbar an den Arbeitsalltag von Lehrerinnen und Lehrern an und vermittelt dadurch die gewünschte Handlungsnahe. Der verwendete SKR^{Bio} in Studie 2 und 3 geht einen Schritt weiter und involviert die Lehramtsstudierenden direkt in das Unterrichtsgeschehen, indem die Rolle der Lehrerin oder des Lehrers übernommen wird. Die freie Auswahl der Fragen durch die Lehrkraft sowie die darauf gegebenen individualisierten Antworten der virtuellen Schülerinnen und Schüler lassen eine dynamische Interaktion aller Akteure im SKR^{Bio} entstehen. Diese sogenannten Rollenspielsimulationen erwiesen sich schon in anderen Studien als ein guter Ansatz, die Lücke zwischen realen Unterrichtssituationen und der theoretischen Lehramtsausbildung zu schließen, sowie Erkenntnisse über handlungsorientiertes Wissen zu generieren (z. B. Hixon und So, 2009). Auf dieser Grundlage kann das gezeigte professionelle Wissen der

DISKUSSION UND AUSBLICK

Lehramtsstudierenden im SKR^{Bio} als prozedurales Wissen operationalisiert werden, da hier das Wissen in einer konkreten Handlungssituation angewendet werden musste, um die gegebenen virtuellen Schülerantworten innerhalb der simulierten Unterrichtsstunde im SKR^{Bio} zu diagnostizieren. Weiterhin kann dieses gezeigte Wissen im SKR^{Bio} analog zu dem Wissen gesehen werden, welches während einer realen Klassendiskussion zum Prozess der natürlichen Selektion benötigt wird. Ein großer Unterschied ist die jeweilige Komplexität der Unterrichtssituation. Während sich in einem realen Klassenraum mehrere Reize akkumulieren, auf welche die Lehrkraft reagieren muss, bietet eine simulierte Klassenraumumgebung durch kontrollierbare Variablen eine deutlich komplexitätsreduzierte Lernumgebung (z. B. Cheong, 2010). Diese augenscheinliche Limitation birgt den Vorteil, dass somit prozedurales Professionswissen isoliert beobachtet werden kann, ohne eine Verzerrung durch äußere Einflüsse berücksichtigen zu müssen.

Zur theoretischen Rahmung des Professionswissens in den durchgeführten Studien wurden zwei Modelle genutzt. Zur Validierung der Testergebnisse im SKR^{Bio} hinsichtlich des prozeduralen FDW in Studie 2 wurde das *RCM* verwendet. Favorisiert wurde dieses aus der Fachdidaktik stammende Modell, da es vor allem verschiedene Ausprägungen des FDW fokussiert (Carlson und Daehler, 2019). Entsprechend der Konzeptualisierung im *RCM* wurde in Studie 2 überprüft, ob der SKR^{Bio} valide Ergebnisse zum ausgeführten FDW (engl. enacted *PCK*) der Lehramtsstudierenden in der Facette Wissen über Schülerfehlvorstellungen liefert. Analog zu der kognitionspsychologischen Kategorisierung von Wissenstypen kann das *ePCK* dem prozeduralen FDW zugeordnet werden. Da in den Studien 1 und 3 der Fokus nicht ausschließlich auf dem FDW lag, wurde hier zur Konzeptualisierung der Wissensbereiche auf das *COACTIV*-Modell zurückgegriffen (Baumert und Kunter, 2006), wobei das FW, FDW und PW weiter in deklaratives und prozedurales Wissen differenziert wurden.

Besonders in Studie 3 war diese theoretische Rahmung durch das *COACTIV*-Modell notwendig, da hier zusätzlich zum SKR^{Bio} ein Fragebogen genutzt wurde, um neben den prozeduralen auch deklarative Facetten des professionellen Wissens der Lehramtsstudierenden zu erfassen. Hier kann das *RCM* nur schwer adaptiert werden, da es keine expliziten Aussagen zum prozeduralen FW und PW liefert. Das Wissen, welches benötigt wurde, um die Fragen im Fragebogen zu beantworten, wurde nach kognitionspsychologischen Ansätzen als deklaratives Wissen definiert, da hier vor allem Wissen über Fakten, Begriffe und Prinzipien zum Tragen kommt (z. B. Krathwohl, 2002). Die Ergebnisse der Studie 3 zeigten, dass das deklarative und prozedurale Professionswissen angehender Biologielehrkräfte zusammenhängen. Gleichzeitig werden theoretische sowie empirische Erkenntnisse aus der kognitionspsychologischen Forschung repliziert (z. B. Shavelson et al., 2005).

8.1.3 Erfassung des prozeduralen FDW mit dem SKR^{Bio}

In allen drei Studien wurden verschiedene Facetten des Professionswissens zur Evolution angehender Lehrkräfte untersucht, wobei der Fokus auf der Erfassung des biologiebezogenen FDW in der Facette Wissen über Schülerfehlvorstellungen lag. Hierbei wurde innerhalb der Studien das FDW über die Fähigkeit der angehenden Lehrerinnen und Lehrer operationalisiert, eine spezifische Fehlvorstellungskategorie in den virtuellen Schülertexten im SI (Studie 1) bzw. den virtuellen Schülerantworten im SKR^{Bio} (Studie 2 und 3) zu diagnostizieren. Diese Konzeptualisierung folgt der Annahme, dass das Wissen über Schülerfehlvorstellungen maßgeblich dafür verantwortlich ist, dieselben auch bei Schülerinnen und Schülern zu erkennen (z. B. Förtsch et al., 2018).

Die Ergebnisse der durchgeführten Studien zeigten, dass die angehenden Lehrerinnen und Lehrer in der Lage waren, wissenschaftlich korrekte von fehlerhaften Denkweisen zur natürlichen Selektion in den virtuellen Schülertexten bzw. Schülerantworten zu differenzieren. Um diese Unterscheidung vornehmen zu können, wird FW in der Facette Wissen über Evolutionsmechanismen benötigt (z. B. Großschedl et al., 2015b). Deutlich problematischer schien die Diagnose der spezifischen Fehlvorstellungskategorien zu sein. Über alle drei Studien hinweg wurden im Mittel nur etwa die Hälfte der artikulierten spezifischen Fehlvorstellungskategorien innerhalb der virtuellen Schülertexte bzw. Schülerantworten durch die angehenden Lehrerinnen und Lehrer erkannt. Eine vergleichende Betrachtung der Anteile korrekt diagnostizierter spezifischer Fehlvorstellungskategorien in den drei durchgeführten Studien zeigte sehr ähnliche Befunde. Diese Replikation der Ergebnisse in den Studien weist darauf hin, dass die beobachteten Schwierigkeiten mit den anthropomorphen, lamarckistischen und teleologischen Fehlvorstellungen generalisierbar sind. Im Durchschnitt wurde am häufigsten die lamarckistische Fehlvorstellung erkannt, gefolgt von anthropomorphen Äußerungen der virtuellen Schülerinnen und Schüler. Die größten Probleme hatten die angehenden Lehrerinnen und Lehrer, wenn es um die Diagnose der teleologischen Fehlvorstellung ging. In Studie 1 konnte darüber hinaus beobachtet werden, dass Schülertexte, welche eine teleologische Fehlvorstellung beinhalteten, besser bewertet wurden als Schülertexte mit einer anthropomorphen Fehlvorstellung. Diese Befunde sind in vielerlei Hinsicht kritisch. Die Hauptaufgabe einer Lehrkraft im Biologieunterricht ist es, die Schülerinnen und Schüler beim Erlernen des wissenschaftlich korrekten Konzepts zu unterstützen und fehlerhafte Aussagen oder Erklärungen zu erkennen, um direkt darauf reagieren zu können (z. B. Gregory, 2009). Hierfür muss innerhalb der akademischen Laufbahn ein elaboriertes Professionswissen erworben werden, um den Anforderungen in realen Unterrichtssituationen gewachsen zu sein. Die durchgeführten Studien legen nahe, dass vor allem das prozedurale FDW in der Facette Wissen über Schülerfehlvorstellungen bei den untersuchten Lehramtsstudierenden bzw. LiV der Biologie nicht adäquat

angewendet werden konnte. Diese Befunde schließen an die bisherige biologiedidaktische Fehlvorstellungsforschung an, wonach Personen aller Altersgruppen und unterschiedlichster Bildungsgrade Schwierigkeiten mit der Evolutionstheorie haben (z. B. Nehm und Schonfeld, 2007). Besonders brisant ist der Befund, dass angehende Lehrkräfte hinsichtlich der Bewertung von Leistungen zwischen verschiedenen Fehlvorstellungskategorien zu teilweise falschen Ergebnissen kommen. So wird beispielsweise die anthropomorphe Fehlvorstellung als vermeintlich unwissenschaftlicher eingeschätzt als die teleologische Fehlvorstellung. Auch wurde die teleologische Fehlvorstellung nicht so häufig als fehlerhafte Denkweise diagnostiziert, wie dies bei der anthropomorphen oder lamarckistischen Fehlvorstellung der Fall war. Eine Erklärung hierfür könnten empirische Studien geben, welche zeigten, dass viele alltägliche Prozesse ziel- bzw. zweckorientiert erklärt werden. Kommt es zu einem Transfer dieser Erklärungsschemata auf den evolutionären Prozess der natürlichen Selektion, können Fehlvorstellungen daraus resultieren (Gresch und Martens, 2019). Weiterhin gibt es Hinweise darauf, dass Lehrerinnen und Lehrer, welche über eine Fehlvorstellung zur Evolution verfügen, nicht nur Probleme bei der Diagnose dieser spezifischen Fehlvorstellung in Schülerantworten haben, sondern diese sogar im Unterricht fördern können (z. B. Yates und Mareck, 2014).

Vor dem Hintergrund dieser Befunde muss sich die universitäre Lehre bezüglich der zahlreichen Fehlvorstellungen zur Evolution in den Köpfen angehender Lehrerinnen und Lehrer bewusst sein. Darauf aufbauend müssen adäquate Lehr- und Lernkonzepte erarbeitet werden, um diesen Fehlvorstellungen zu begegnen, und möglichst die Entwicklung wissenschaftlich korrekter Denkweisen zu fördern. Unterstützend könnte hierbei der SKR^{Bio} sein, um vorerst zu eruieren, welche Fehlvorstellungen von den angehenden Lehrkräften nicht diagnostiziert werden können bzw. ob es einen expliziten Schwerpunkt bei einer spezifischen Fehlvorstellungskategorie gibt. Die Befunde im SKR^{Bio} würden entsprechend der Validierung in Studie 2 gleichzeitig direkte Rückschlüsse auf das prozedurale FDW in der Facette Wissen zu Schülerfehlvorstellungen zulassen.

8.2 Grenzen der Studien

Die im Rahmen dieser Dissertation angefertigten Manuskripte setzen sich bereits differenziert mit verschiedenen Grenzen der durchgeführten Studien auseinander (Kapitel 4 bis 6). Im Folgenden werden die relevantesten Grenzen thematisiert.

Inhaltliche Entwicklung des SKR^{Bio}. Ziel der inhaltlichen Entwicklung des SKR^{Bio} war es, neben dem PW auch weitere Facetten des evolutionsbezogenen professionellen Wissens angehender Lehrerinnen und Lehrer zu erfassen. Dementsprechend wurden

evolutionsbiologische Fragestellungen entwickelt, die zwischen verschiedenen Organismen und Merkmalsausprägungen variierten. Zu diesen Fragen wurden virtuelle Schülerantworten konzipiert, welche eine spezifische Fehlvorstellung beinhalten. Diese spezifischen Fehlvorstellungskategorien basierten auf bisheriger biologiedidaktischer Forschung, wobei in den durchgeführten Studien anthropomorphe, lamarckistische und teleologische Fehlvorstellungen fokussiert wurden. Obwohl diese Fehlvorstellungen einen großen Anteil aller geäußerten Fehlvorstellungen ausmachen, gibt es in der Forschung Hinweise auf weit mehr fehlerhafte Denkweisen zur Evolutionstheorie, die nicht im SKR^{Bio} adressiert wurden (z. B. Graf und Hamdorf, 2011). Weiterhin werden in Erklärungen zur natürlichen Selektion häufig mehrere Fehlvorstellungen miteinander kombiniert (z. B. Zohar und Ginossar, 1998). Dieses beobachtete Phänomen wurde im SKR^{Bio} nicht berücksichtigt, um vorerst die Komplexität möglichst gering zu halten. Die durchgeführten Studien konzentrierten sich vor allem auf das FDW zum Prozess der natürlichen Selektion, der zu Anpassungen innerhalb von Populationen führt. Andere evolutionsbiologische Prozesse, wie der Genfluss oder der genetische Drift, wurden bei der Entwicklung des SKR^{Bio} vorerst vernachlässigt (z. B. Beggrow und Nehm, 2012).

Erfassung des Professionswissens. Der SKR^{Bio} ermöglicht die Erfassung verschiedener Facetten des evolutionsbezogenen Professionswissens, wobei in den Studien explizit die Facetten: Wissen über Evolutionsmechanismen (FW), Wissen über Schülerfehlvorstellungen (FDW) und das pädagogische Beurteilungswissen (PW) adressiert wurden. Die Professionsforschung identifizierte innerhalb des FW, FDW und PW zahlreiche weitere relevante Wissensfacetten, die Lehrerinnen und Lehrer benötigen, um erfolgreich und effektiv Unterricht zu gestalten. Beispielsweise differenziert das FDW zusätzlich in die Facette des Wissens über Instruktionsstrategien. Das wird besonders nach der Diagnose einer spezifischen Fehlvorstellungskategorie relevant, um lernförderliche Maßnahmen einzuleiten (z. B. Magnusson et al., 1999). Diese und weitere in der Forschung beschriebene Facetten des Professionswissens können derzeit nicht mit dem SKR^{Bio} erfasst werden.

Erfassung prozeduralen Wissens. Der SKR^{Bio} wurde dahingehend entwickelt, dass Facetten prozeduralen Professionswissens zur Evolution erfasst werden konnten. Grundlage hierfür war die aus der kognitionspsychologischen Forschung stammende Differenzierung in deklaratives und prozedurales Wissen, wonach das gezeigte Professionswissen der Lehramtsstudierenden im SKR^{Bio} operationalisiert und einer der beiden Wissenstypen zugeordnet werden konnte (z. B. Anderson und Lebière, 1998; König et al., 2014). Die Betrachtung der tatsächlich ablaufenden sensorischen Gedächtnisleistungen, wie es zum Beispiel die ACT Theorie (engl. *Adaptive Control of Thought Theory*) veranschaulicht (Anderson, 1983), wurde dabei nicht weiter berücksichtigt. Dementsprechend wurde im SKR^{Bio} das direkt in der Handlung gezeigte

Wissen (Handlungswissen), das aus sensorischen Gedächtnisleistungen resultierte, als prozedurales Wissen operationalisiert und gemessen.

Generalisierbarkeit der Ergebnisse. Die Generalisierbarkeit der im Rahmen dieser Dissertation gewonnenen Befunde muss durch zusätzliche Studien, möglichst auf nationaler sowie internationaler Ebene, repliziert werden. Obwohl einige Ergebnisse, wie beispielsweise die Diagnoseleistung von angehenden Lehrerinnen und Lehrern hinsichtlich spezifischer Fehlvorstellungskategorien repliziert werden konnten, sind weitere Untersuchungen mit diversen Stichproben (bspw. Nationalität, Studienfortschritt, Schulform) wünschenswert.

8.3 Implikationen für zukünftige Forschung und universitäre Praxis

In den vergangenen Jahren wurde einige simulierte Klassenraumumgebungen entwickelt, um prozedurale Facetten des PW zu messen und zu fördern (z. B. Bautista und Boon, 2015; Dalgarno et al., 2016). Der im Rahmen dieser Dissertation entwickelte und erprobte SKR^{Bio} ermöglicht darüber hinaus die Erfassung von verschiedenen Facetten prozeduralen FW und FDW zur Evolution. Durch das adaptive Design dieser simulierten Klassenraumumgebung könnten unterschiedliche experimentelle Variationen untersucht werden. Denkbar wäre, den bestehenden SKR^{Bio} für andere Fachbereiche umzustrukturieren, um beispielsweise das Wissen über Schülerfehlvorstellungen zum Energiekonzept in der Physik oder Chemie zu untersuchen. Voraussetzung hierfür wäre die Konzeption neuer Fragestellungen und möglicher virtueller Schülerantworten, welche spezifisch die neue Thematik erfassen und zu den jeweiligen Forschungsfragen passen.

Auf Basis des gezeigten FDW der angehenden Lehrerinnen und Lehrer in den durchgeführten Studien zeigt sich, dass es weiterhin Forschungsbedarf hinsichtlich spezifischer Fehlvorstellungskategorien zum Prozess der natürlichen Selektion gibt. Nach wie vor scheinen fehlerhafte Denkweisen bei den Lehramtsstudierenden zu existieren, die sich in Zukunft negativ auf die diagnostischen Aktivitäten im Biologieunterricht auswirken könnten. Dementsprechend muss die universitäre Lehramtsausbildung konsequent die Probleme beim Lehren und Lernen der Evolution adressieren, um bei angehenden Lehrerinnen und Lehrern ein Bewusstsein für die zahlreichen Fehlvorstellungen zu schaffen. Weiterhin müsste empirisch überprüft werden, ob andere Fehlvorstellungskategorien, wie zum Beispiel der Essentialismus, zu ähnlichen Schwierigkeiten bei der Beurteilung von Schülertexten bzw. –antworten führen. Obwohl viele empirische Studien über spezifische Fehlvorstellungskategorien zur Evolution vorhanden sind (z. B. Gregory, 2009), gibt es in der aktuellen biologiedidaktischen Forschung kaum Erkenntnisse darüber, inwiefern diagnostizierte Fehlvorstellungen weiter

genutzt werden können, um darauf aufbauend oder durch Konfrontation korrekte Denkweisen zu initiieren. Relevant wird in diesem Zusammenhang das Wissen über Instruktionsstrategien (z. B. Harms und Reiss, 2019; Ziadie und Andrews, 2018). Dementsprechend wird weitere Forschung benötigt, um eine begründete Auswahl von Anschlusshandlungen auf Basis der diagnostizierten Fehlvorstellung empirisch abzusichern.

Die Ergebnisse der Studie 1 und Studie 3 lieferten erste Hinweise darauf, dass ein Zusammenhang zwischen dem deklarativen und dem prozeduralen biologiebezogenen Wissen besteht. Damit konnte an bisherige Erkenntnisse aus der kognitionspsychologischen Forschung angeknüpft werden, wonach beide Wissenstypen miteinander interagieren (z. B. Shavelson et al., 2005). Trotzdem sind die Befunde im SKR^{Bio} nicht für andere Fachbereiche oder Professionswissensfacetten übertragbar. Erst weitere fachdidaktische Studien, welche die Struktur sowie die Interaktion deklarativen und prozeduralen Wissens empirisch untersuchen, könnten zu generalisierbaren Ergebnissen in diesem Forschungsdesiderat beitragen. Der SKR^{Bio} scheint den Lehramtsstudierenden eine adäquate digitale Lernumgebung zu bieten, das im akademischen Kontext erworbene deklarative biologiebezogene Professionswissen in einer konkreten Handlungssituation anzuwenden. Besonders relevant ist diese Erkenntnis vor dem Hintergrund, dass deklaratives Wissen durch kontinuierliche Übung und Kontextualisierung in prozedurales Wissen transformiert werden kann (z. B. Anderson, 1982; Schneider und Stern, 2010). Folglich sollten den Lehramtsstudierenden bereits während der universitären Ausbildung möglichst viele Gelegenheiten gegeben werden, mit simulierten Klassenraumumgebungen, wie beispielsweise dem SKR^{Bio}, zu arbeiten, um die Prozeduralisierung des Wissens bereits während des Studiums zu fördern.

8.4 Ausblick

Bereits seit einigen Jahren beschäftigt sich die naturwissenschaftsdidaktische Forschung mit neuen digitalen Möglichkeiten, um deren Potenziale für wissenschaftliche Zwecke oder eine effektivere Lehre einzusetzen. Zeitgleich gibt es den politischen Willen, digitale Projekte im Bereich der Hochschulbildung zu fördern. Das vorliegende Dissertationsprojekt wurde im Rahmen der ersten Förderlinie – Forschung zur digitalen Hochschulbildung – des Bundesministeriums für Bildung und Forschung (BMBF) initiiert (Wissenschafts- und Hochschulforschung, 2021). Nachdem der SKR^{Bio} als Messinstrument entwickelt, validiert und erprobt wurde, wird dieser in den kommenden Jahren als Trainingsinstrument weiterentwickelt. Nach erfolgreicher Antragsstellung beim BMBF in der vierten Förderlinie – Innovationen in der Hochschulbildung durch Künstliche Intelligenz und Big Data – werden zukünftig Feedbackbots (d.h. automatisierte, technische

DISKUSSION UND AUSBLICK

Dialogsysteme) in den SKR^{Bio} implementiert. Dieses Vorhaben hat das Ziel, mithilfe künstlich-intelligenter Feedbackbots eine adaptive Rückmeldung zu den gezeigten Leistungen der Nutzerinnen und Nutzer im SKR^{Bio} zu geben, um somit die Entwicklung von prozeduralem biologiebezogenen Professionswissen nicht nur messen, sondern auch fördern zu können.

LITERATURVERZEICHNIS

Anmerkung: Die Literaturangaben zu den einzelnen Manuskripten (Studie 1 bis 3) sind jeweils am Ende der Kapitel 4, 5 und 6.

- Alexander, P.A., & Judy, J.E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research*, 58, 375–405. <https://doi.org/10.3102/00346543058004375>
- Alters, B. J., & Nelson C. E. (2002). Perspective: Teaching evolution in higher education. *Evolution: International Journal of Organic Evolution*, 56(10), 1891–1901. <https://doi.org/10.1111/j.0014-3820.2002.tb00115.x>
- Anderson, J. R. (1976). *Language, memory and thought*. Hillsday, NJ: Erlbaum.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge: Harvard University Press.
- Anderson, J. R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum Akademischer Verlag.
- Anderson, J. R., & Lebière, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Andersson, B., & Wallin, A. (2006). On developing content-oriented theories taking biological evolution as an example. *International Journal of Science Education*, 28(6), 673–695. <https://doi.org/10.1080/09500690500498385>
- Andrews, T. M., Kalinowski, S. T., & Leonard, M. J. (2011). "Are humans evolving?" A classroom discussion to change student misconceptions regarding natural selection. *Evolution: Education and Outreach*, 4, 456–466. doi: 10.1007/s12052-011-0343-4
- Baalmann, W., Frerichs, V., Weitzel, H., Gropengießer, H., & Kattmann, U. (2004). Schülervorstellungen zu Prozessen der Anpassung–Ergebnisse einer Interviewstudie im Rahmen der Didaktischen Rekonstruktion. *Zeitschrift für Didaktik der Naturwissenschaften*, 10(1), 7–28.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Badiee, F., & Kaufman, D. (2015). Design evaluation of a simulation for teacher education. *Sage Open*, 5(2), 1–10. <https://doi.org/10.1177/2158244015592454>

LITERATURVERZEICHNIS

- Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching. *American Educator*, 29, 14-46.
- Basel, N., Harms, U., Precht, H., Weiss, T., & Rothgangel, M. (2014). Students' arguments on the science and religion issue: the example of evolutionary theory and genesis. *Journal of Biological Education*, 48(4), 179–187. <https://doi.org/10.1080/00219266.2013.849286>
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Baumert, J. & Kunter, M. (2011a). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.
- Baumert, J., & Kunter, M. (2011b). Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern 1. *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*, 163–192.
- Baumert, J., & Kunter, M. (2013). Motivation as an aspect of professional competence: Research findings on teacher enthusiasm. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (S. 273-289). New York: Springer.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133–180. <https://doi.org/10.3102/0002831209345157>
- Bautista, N. U., & Boone, W. J. (2015). Exploring the impact of TeachME™ Lab virtual classroom teaching simulation on early childhood education majors' self-efficacy beliefs. *Journal of Science Teacher Education*, 26(3), 237–262. <https://doi.org/10.1007/s10972-014-9418-8>
- Beardsley, P. M. (2004). Middle school student learning in evolution: Are current standards achievable?. *The American Biology Teacher*, 66(9), 604–613.
- Beggrow, E. P., & Nehm, R. H. (2012). Students' mental models of evolutionary causation: natural selection and genetic drift. *Evolution: Education and Outreach*, 5(3), 429–444. doi:10.1007/s12052-012-0432-z

- Berry, A., Friedrichsen, P., & Loughran, J. (Hrsg.). (2015). *Re-examining pedagogical content knowledge in science education*. New York: Routledge. <https://doi.org/10.4324/9781315735665>
- Besser, M., & Krauss, S. (2009). Zur Professionalität als Expertise. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrprofessionalität - Bedingungen, Genese, Wirkungen und ihre Messung*. Weinheim und Basel: Beltz.
- Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of research in science teaching*, 27(5), 415–427. <https://doi.org/10.1002/tea.3660320509>
- Bizzo, N. M. V. (1994). From down house landlord to Brazilian high school students: what has happened to evolutionary knowledge on the way? *Journal of Research in Science Teaching*, 31(5), 537–556. <https://doi.org/10.1002/tea.3660310508>
- Blömeke, S., Felbrich, A., & Müller, C. (2008). Messung des erziehungswissenschaftlichen Wissens angehender Lehrkräfte. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare* (S. 171–218). Münster: Waxmann.
- Blömeke, S., Kaiser, G., Döhrmann, M., & Lehmann, R. (2010). Mathematisches und mathematikdidaktisches Wissen angehender Sekundarstufen-I-Lehrkräfte im internationalen Vergleich. In S. Blömeke, G. Kaiser, & R. Lehmann (Eds.), *TEDS-M 2008: Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich* (S. 197–238). Münster: Waxmann.
- Blömeke, S., Kaiser, G., Lehmann, R., König, J., Döhrmann, M., Buchholtz, C., & Hacke, S. (2009). TEDS-M: Messung von Lehrerkompetenzen im internationalen Vergleich. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 181–210). Weinheim: Beltz.
- Borowski, A., Kirschner, S., Liedtke, S. & Fischer, H. E. (2011). Vergleich des Fachwissens von Studierenden, Referendaren und Lehrenden in der Physik. *PhyDid- Physik und Didaktik in Schule und Hochschule.*, 10(1), 1–9.
- Borowski, A., Neuhaus, B. J., Tepner, O., Wirth, J., Fischer, H. E., Leutner, D., Sandmann, A., & Sumfleth, E. (2010). Professionswissen von Lehrkräften in den Naturwissenschaften (ProwiN)–Kurzdarstellung des BMBF-Projekts. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 341–349.

- BouJaoude, S., Asghar, A., Wiles, J. R., Jaber, L., Saredidine, D., & Alters, B. (2011). Biology professors' and teachers' positions regarding biological evolution and evolution education in a Middle Eastern society. *International Journal of Science Education*, 33(7), 979–1000. <https://doi.org/10.1080/09500693.2010.489124>
- Bradley, E. G., & Kendall, B. (2014). A review of computer simulations in teacher education. *Journal of Educational Technology Systems*, 43(1), 3–12. <https://doi.org/10.2190/ET.43.1.b>
- Bravo, P., & Cofré, H. (2016). Developing biology teachers' pedagogical content knowledge through learning study: the case of teaching human evolution. *International Journal of Science Education*, 38(16), 2500–2527. <https://doi.org/10.1080/09500693.2016.1249983>
- Bromme, R. (1992). *Der Lehrer als Experte. Zur Psychologie des professionellen Lehrerwissens [The expert teacher. Psychology of professional knowledge]*. Bern: Verlag Hans Huber.
- Bromme, R. (1994). Beyond subject matter: A psychological topology of teachers' professional knowledge. *Didactics of mathematics as a scientific discipline*, 73–88.
- Bromme, R. (2001). Teacher Expertise. In N. J. Smelser, P. B. Baltes, & F. E. Weinert (Hrsg.), *International Encyclopedia of the Behavioral Sciences: Education* (S. 15459–15465). London: Pergamon.
- Bruckermann, T., Fiedler, D. & Harms, U. (2021). Identifying precursory concepts in evolution during early childhood: A systematic literature review. *Studies in Science Education*, 57(1), 85–127. doi.org/10.1080/03057267.2020.1792678
- Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education*, 68(4), 493–503. <https://doi.org/10.1002/sce.3730680412>
- Byrne, J., Grace, M., & Hanley, P. (2009). Children's anthropomorphic and anthropocentric ideas about micro-organisms: Educational research. *Journal of Biological Education*, 44(1), 37–43. <https://doi.org/10.1080/00219266.2009.9656190>
- Carlsen, W. (1999). Domains of teacher knowledge. In J. Gess-Newsome & N. G. Lederman (Hrsg.), *Examining pedagogical content knowledge* (S. 133–144). Dordrecht: Kluwer Academic Publishers.
- Carlson, J., & Daehler, K. R. (2019). The refined consensus model of pedagogical content knowledge in science education. In *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (S. 77–92). Springer. https://doi.org/10.1007/978-981-13-5898-2_2

- Chan, K. K. H., & Hume, A. (2019). Towards a consensus model: Literature review of how science teachers' pedagogical content knowledge is investigated in empirical studies. *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science*, 3–76. https://doi.org/10.1007/978-981-13-5898-2_1
- Cheong, D. (2010). The effects of practice teaching sessions in Second Life on the change in pre-service teachers' teaching efficacy. *Computers and Education*, 55(2), 868–880. <https://doi.org/10.1016/j.compedu.2010.03.018>
- Christensen, R., Knezek, G., & Tyler-Wood, T. (2011). simSchool: An online dynamic simulator for enhancing teacher preparation. *International Journal of Learning Technology*, 6(2), 201–220. <https://doi.org/10.1504/IJLT.2011.042649>
- Clapper, T. C. (2010). Role play and simulation returning to teaching for understanding. *The Education Digest*, 75(8), 39–43.
- Cochran, K. F., DeRuiter, J. A., & King, R. A. (1993). Pedagogical content knowing: An integrative model for teacher preparation. *Journal of Teacher Education*, 44(4), 263–272. <https://doi.org/10.1177/0022487193044004004>
- Coley, J. D., & Tanner, K. D. (2012). Common origins of diverse misconceptions: cognitive principles and the development of biology thinking. *CBE-Life Sciences Education*, 11, 209–215. doi:10.1187/cbe.12-06-0074
- Dagher, Z. R., & BouJaoude, S. (1997). Scientific views and religious beliefs of college students: The case of biological evolution. *Journal of Research in Science Teaching*, 34(5), 429–445. [https://doi.org/10.1002/\(SICI\)1098-2736\(199705\)34:5<429::AID-TEA2>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1098-2736(199705)34:5<429::AID-TEA2>3.0.CO;2-S)
- Dalgarno, B., Bishop, A. G., Adlong, W., & Bedgood Jr, D. R. (2009). Effectiveness of a virtual laboratory as a preparatory resource for distance education chemistry students. *Computers & Education*, 53(3), 853–865. <https://doi.org/10.1016/j.compedu.2009.05.005>
- Dalgarno, B., Gregory, S., Knox, V., & Reiners, T. (2016). Practising teaching using virtual classroom role plays. *Australian Journal of Teacher Education*, 41(1), 126–154. <https://doi.org/10.14221/ajte.2016v41n1.8>
- Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323(5910), 66–69. <https://doi.org/10.1126/science.1167311>
- De Jong, T. & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31, 105–113.

- De Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179–201. doi:10.3102/00346543068002179
- Demastes, S. S., Good, R. G., & Peebles, P. (1996). Patterns of conceptual change in evolution. *Journal of Research in Science Teaching*, 33(4), 407–431. [https://doi.org/10.1002/\(SICI\)1098-2736\(199604\)33:4<407::AIDTEA4>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1098-2736(199604)33:4<407::AIDTEA4>3.0.CO;2-W)
- Demastes, S. S., Settlage Jr, J., & Good, R. (1995). Students' conceptions of natural selection and its role in evolution: Cases of replication and comparison. *Journal of Research in Science Teaching*, 32(5), 535-550. <https://doi.org/10.1002/tea.3660320509>
- Driver, R. (1988). Changing conceptions. *Journal of Research in Education*, 161-196.
- Edelmann, W. (2000). *Lernpsychologie*. Kempten: Kösel-Verlag.
- European Commission. (2015). *Strengthening teaching in Europe. New evidence from teachers compiled by Eurydice and CRELL*. Retrieved from http://ec.europa.eu/education/library/policy/teaching-profession-practices_en.pdf.
- Evans, E. M. (2000). The emergence of beliefs about the origins of species in school-age children. *Merrill-Palmer Quarterly* (1982), 221–254.
- Fenstermacher, G. (1994). The knower and the known. The nature of knowledge in research on teaching. In L. Darling-Hammond (Hrsg.), *Review of research education* (Bd. 20, S. 3–56). Washington, DC: American Educational Research Association.
- Ferrari, M., & Chi, M. T. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education*, 20(10), 1231–1256. <https://doi.org/10.1080/0950069980201005>
- Ferry, B., Kervin, L., Cambourne, B., Turbill, J., Puglisi, S., Jonassen, D., & Hedberg, J. (2004). Online classroom simulation: The next wave for pre-service teacher education. In *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (S. 294–302). Perth: Australian Society for Computers in Learning in Tertiary Education.
- Fiedler, K., Freytag, P. & Unkelbach, C. (2007). Pseudo-contingencies in a simulated classroom. *Journal of Personality and Social Psychology*, 92, 665–677. <https://doi.org/10.1037/0022-3514.92.4.665>
- Fiedler, D., Tröbst, S., Großschedl, J., & Harms, U. (2018). EvoSketch: Simple simulations for learning random and probabilistic processes in evolution, and effects of

- instructional support on learners' conceptual knowledge. *Evolution: Education and Outreach*, 11(1), 1–17. <https://doi.org/10.1186/s12052-018-0089-3>
- Fiedler, K., Walther, E., Freytag, P. & Plessner, H. (2002). Judgment biases in a simulated classroom – A cognitive-environmental approach. *Organizational Behavior and Human Decision Processes*, 88, 527–561. <https://doi.org/10.1006/obhd.2001.2981>
- Fischer, H. E., Borowski, A., & Tepner, O. (2012). Professional knowledge of science teachers. In *Second International Handbook of Science Education* (S. 435–448). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-9041-7_30
- Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M., Girwidz, R., Obersteiner, A., ... & Seidel, T. (2018). Systematizing professional knowledge of medical doctors and teachers: development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences*, 8(4), 207–225. <https://doi.org/10.3390/educsci8040207>
- Fosnot, C. T., & Perry, R. S. (2005). Constructivism: A psychological theory of learning. In C. T. Fosnot (Hrsg.), *Constructivism: Theory, perspectives, and practice* (2. Auflage, S. 8–33). New York, NY: Teachers College Press.
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, 49(9), 1181–1210. <https://doi.org/10.1002/tea.21054>
- Gess-Newsome, J. (1999). Pedagogical content knowledge: An introduction and orientation. In J. Gess-Newsome & N. G. Lederman (Hrsg.), *Examining pedagogical content knowledge* (S. 3–17). Dordrecht: Kluwer Academic Publishers.
- Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK summit. In A. Berry, P. Friedrichsen, & J. Loughran (Hrsg.), *Re-examining pedagogical content knowledge in science education*. New York, London: Routledge.
- Gess-Newsome, J., & Lederman, N. G. (Hrsg.). (1999). *Explaining pedagogical content knowledge*. Dordrecht: Kluwer Academic.
- Girod, M., & Girod, G. R. (2008). Simulation and the need for practice in teacher preparation. *Journal of Technology and Teacher Education*, 16(3), 307–337.
- Graf, D. & Hamdorf, E. (2011): Evolution: Verbreitete Fehlvorstellungen zu einem zentralen Thema. In: Dreesmann, D., Graf, D. & Witte, K. [Hrsg.]: *Evolutionsbiologie: Moderne Themen für den Unterricht*, (S. 25–41). Heidelberg: Springer Akademischer Verlag.

LITERATURVERZEICHNIS

- Graf, D., & Soran, H. (2010). Einstellung und Wissen von Lehramtsstudierenden zur Evolution—ein Vergleich zwischen Deutschland und der Türkei. In *Evolutionstheorie-Akzeptanz und Vermittlung im europäischen Vergleich* (S. 141–161). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-02228-9_10
- Gramzow, Y., Riese, J. & Reinhold, P. (2013). Modellierung fachdidaktischen Wissens angehender Physiklehrkräfte. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 7–30.
- Gregory, T. R. (2009). Understanding Natural selection. Essential concepts and common misconceptions. *Evolution: Education and Outreach*, 2(2), 156–175. <https://doi.org/10.1007/s12052-009-0128-1>
- Gresch, H., & Martens, M. (2019). Teleology as a tacit dimension of teaching and learning evolution: A sociological approach to classroom interaction in science education. *Journal of Research in Science Teaching*, 56(3), 243–269.
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teacher College Press.
- Großschedl, J., Harms, U., Kleickmann, T., & Glowinski, I. (2015a). Preservice biology teachers' professional knowledge: Structure and learning opportunities. *Journal of Science Teacher Education*, 26(3), 291–318. <https://doi.org/10.1007/s10972-015-9423-6>
- Großschedl, J., Konnemann, C., & Basel, N. (2014). Pre-service biology teachers' acceptance of evolutionary theory and their preference for its teaching. *Evolution: Education and Outreach*, 7(18), 1–16. doi:10.1186/s12052-014-0018-z
- Großschedl, J., Neubrand, C., Kirchner, A., Oppermann, L., Basel, N., & Gantner, S. (2015b). Entwicklung und Validierung eines Testinstruments zur Erfassung des evolutionsbezogenen Professionswissens von Lehramtsstudierenden (ProwiE). *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 173–185.
- Großschedl, J., Seredszus, F. & Harms, U. (2018). Angehende Biologielehrkräfte: evolutionsbezogenes Wissen und Akzeptanz der Evolutionstheorie. *Zeitschrift für Didaktik der Naturwissenschaften*. 24(1), 51–70.
- Gurvitch, R., & Metzler, M. W. (2009). The effects of laboratory-based and field-based practicum experience on pre-service teachers' self-efficacy. *Teaching and Teacher Education*, 25(3), 437–443. <https://doi.org/10.1016/j.tate.2008.08.006>
- Hamman M, & Nehm R. (2020). Teleology and evolution education: introduction to the special issue. *Evolution: Education and Outreach*, 13(16), 1–5.
- Hashweh, M.Z. (2005). Teacher pedagogical constructions: a reconfiguration of pedagogical

- content knowledge. *Teachers and Teaching: Theory and Practice*, 11(3), 273–292. <https://doi.org/10.1186/s12052-020-00130-y>
- Harms, U., & Reiss, M. J. (2019). The present status of evolution education. In U. Harms & M. J. Reiss (Hrsg.). *Evolution Education Re-considered* (S. 1–19). Basel: Springer. <https://doi.org/10.1007/978-3-030-14698-6>
- Hashweh, M.Z. (2005). Teacher pedagogical constructions: a reconfiguration of pedagogical content knowledge. *Teachers and Teaching: Theory and Practice*, 11(3), 273–292. <https://doi.org/10.1080/13450600500105502>
- Hattie, J. (2005). What is the nature of evidence that makes a difference to learning?. *2005-Using data to support learning*, 11–21. <https://doi.org/10.13128/formare-13253>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon: Routledge. <https://doi.org/10.4324/9780203887332>
- Heitz, J. G., Cheetham, J. A., Capes, E. M., & Jeanne, R. L. (2010). Interactive evolution modules promote conceptual change. *Evolution: Education and Outreach*, 3(3), 436–442. <https://doi.org/10.1007/s12052-010-0208-2>
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Klett/Kallmeyer.
- Hill, H.C., Rowan, B., Ball, D. (2005). Effects of Teachers Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal*. 42(2), 371–406. <https://doi.org/10.3102/00028312042002371>
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, 105(1), 11–30. <https://www.jstor.org/stable/10.1086/428763>
- Hixon, E., & So, H. J. (2009). Technology's role in field experiences for preservice teacher training. *Educational Technology & Society*, 12(4), 294–304.
- Jamaludin, A., San Chee, Y., & Ho, C. M. L. (2009). Fostering argumentative knowledge construction through enactive role play in Second Life. *Computers & Education*, 53(2), 317–329. <https://doi.org/10.1016/j.compedu.2009.02.009>
- Jansen, T., Vögelin, C., Machts, N., Keller, S., Köller, O., & Möller, J. (2021). Judgment accuracy in experienced versus student teachers: Assessing essays in English as a foreign language. *Teaching and Teacher Education*, 97, 1–11. <https://doi.org/10.1016/j.tate.2020.103216>
- Jansen, T., Vögelin, C., Machts, N., Keller, S., & Möller, J. (2019). Empirische Arbeit: Das Schülerinventar ASSET zur Beurteilung von Schülerarbeiten im Fach Englisch. Drei experimentelle Studien zu Effekten der Textqualität und der Schülernamen.

- Psychologie in Erziehung und Unterricht*, 66(4), 303-315. <http://dx.doi.org/10.2378/peu2019.art21d>
- Jensen, M. S., & Finley, F. N. (1996). Changes in students' understanding of evolution resulting from different curricular and instructional strategies. *Journal of Research in Science Teaching*, 33(8), 879–900. [https://doi.org/10.1002/\(SICI\)1098-2736\(199610\)33:8<879::AID-TEA4>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1098-2736(199610)33:8<879::AID-TEA4>3.0.CO;2-T)
- Jiménez-Aleixandre, M. P. (1992). Thinking about theories or thinking with theories? A classroom study with natural selection. *International Journal of Science Education*, 14(1), 51–61. <https://doi.org/10.1080/0950069920140106>
- Jungwirth, E. (1977). Should natural phenomena be described teleologically or anthropomorphically?—a science educator's view. *Journal of Biological Education*, 11(3), 191–196. <https://doi.org/10.1080/00219266.1977.9654143>
- Jüttner, M., Boone, W., Park, S., & Neuhaus, B. J. (2013). Development and use of a test instrument to measure biology teachers' content knowledge (CK) and pedagogical content knowledge (PCK). *Educational Assessment, Evaluation and Accountability*, 25, 45–67. <https://doi.org/10.1007/s11092-013-9157-y>
- Jüttner, M., & Neuhaus, B. J. (2013). Das Professionswissen von Biologielehrkräften – Ein Vergleich zwischen Biologielehrkräften, Biologen und Pädagogen. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 31–49. <https://doi.org/10.11114/jets.v1i2.126>
- Kaiser, J., Möller, J., Helm, F., & Kunter, M. (2015). Das Schülerinventar: Welche Schülermerkmale die Leistungsurteile von Lehrkräften beeinflussen. *Zeitschrift für Erziehungswissenschaft*, 18(2), 279–302. <https://doi.org/10.1007/s11618-015-0619-5>
- Kaiser, J., Südkamp, A., & Möller, J. (2016). The effects of student characteristics on teachers' judgment accuracy: Disentangling ethnicity, minority status, and achievement. *Journal of Educational Psychology*, 109, 881–888. <https://doi.org/10.1037/edu0000156>
- Kalinowski, S. T., Leonard, M. J., & Taper, M. L. (2016). Development and validation of the conceptual assessment of natural selection (CANS). *CBE—Life Sciences Education*, 15(4). <https://doi.org/10.1187/cbe.15-06-0134>
- Kallery, M., & Psillos, D. (2004). Anthropomorphism and animism in early years science: Why teachers use them, how they conceptualise them and what are their views on their use. *Research in Science Education*, 34(3), 291–311. <https://doi.org/10.1023/B:RISE.0000044613.64634.03>

- Kampourakis, K., Pavlidi, V., Papadopoulou, M., & Palaiokrassa, E. (2012). Children's teleological intuitions: What kind of explanations do 7-8 year olds give for the features of organisms, artefacts and natural objects? *Research of Science Education*, 42(4), 651-671. doi:10.1007/s11165-011-9219-4
- Kampourakis, K., & Zogza, V. (2008). Students' intuitive explanations of the causes of homologies and adaptations. *Science & Education*, 17(1), 27-47. <https://doi.org/10.1007/s11191-007-9075-9>
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73. <https://doi.org/10.1111/jedm.12000>
- King, C. J. H. (2010). An analysis of misconceptions in science textbooks: Earth science in England and Wales. *International Journal of Science Education*, 32(5), 565-601. <https://doi.org/10.1080/09500690902721681>
- Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften* (Vol. 161). Berlin: Logos Verlag.
- Kirschner, S., Sczudlek, M., Tepner, O., Borowski, A., Fischer, H. E., Lenske, G., ... & Wirth, J. (2017). Professionswissen in den Naturwissenschaften (ProwiN). In *Entwicklung von Professionalität pädagogischen personals* (S. 113-130). Wiesbaden: Springer.
- Kleickmann, T., Großschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V. & Zimmermann, F. (2014). Professional knowledge of student teachers of mathematics and science subjects - Test development within the project KiL. *Unterrichtswissenschaft*, 42(3), 280-288. <https://doi.org/09201403280>
- Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52(6), 876-903.
- KMK [Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland] (2005). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss—Beschluss vom 16.12.2004*. München: Luchterhand.
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A., & Kaiser, G. (2014). Ist das allgemeine pädagogische Wissen von Lehrkräften eine Voraussetzung für das Wahrnehmen und Interpretieren von Unterrichtssituationen? Ein videobasierter Erhebungsansatz. *Teaching and Teacher Education*, 38, 76-88.

LITERATURVERZEICHNIS

- König, J., Blömeke, S., Paine, L., Schmidt, W. H., & Hsieh, F.-J. (2011). General pedagogical knowledge of future middle school teachers: On the complex ecology of teacher education in the United States, Germany, and Taiwan. *Journal of Teacher Education*, 62, 188–201. <https://doi.org/10.1177/0022487110388664>
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M., & Jordan, A. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie, *Journal für Mathematik-Didaktik* (29), 223–258. <https://doi.org/10.3929/ethz-b-000244751>
- Kröger, J., Petersen, S. & Neumann, K. (2013). Messung professioneller Kompetenz im Fach Physik. In S. Bernholt (Hrsg.), *Inquiry-based Learning – Forschendes Lernen* (S. 533–535). Kiel: IPN.
- Kunter, M., Kleickmann, T., Klusmann, U. & Richter, D. (2011). Die Entwicklung professioneller Kompetenz von Lehrkräften. *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*, 1, 55–68.
- Kunter, M. & Pohlmann, B. (2009). Lehrer. In E. Wild & J. Möller (Eds.), *Einführung in die Pädagogische Psychologie* (S. 261–282). Heidelberg: Springer Medizin Verlag.
- Lange, K., Kleickmann, T., Tröbst, S. & Möller, K. (2012). Fachdidaktisches Wissen von Lehrkräften und multiple Ziele im naturwissenschaftlichen Sachunterricht. *Zeitschrift für Erziehungswissenschaft*, 15 (1), 55–75. doi: 10.1007/s11618-012-0258-z
- Lee, E., & Luft, J. A. (2008). Experienced secondary science teachers' representation of pedagogical content knowledge. *International Journal of Science Education*, 30, 1343–1363. <https://doi.org/10.1080/09500690802187058>
- Leonard, M. J., Kalinowski, S. T., & Andrews, T. C. (2014). Misconceptions yesterday, today, and tomorrow. *CBE-Life Sciences Education*, 13, 179–186. doi:10.1187/cbe.13-12-0244
- Loughran, J., Berry, A., & Mulhall, P. (2006). *Understanding and developing science teachers' pedagogical content knowledge*. Rotterdam: Sense.
- Lucero, M. M., Delgado, C., & Green, K. (2019). Elucidating High School Biology Teachers' Knowledge of Students' Conceptions Regarding Natural Selection. *International Journal of Science and Mathematics Education*, 1–21. <https://doi.org/10.1002/tea.21344>
- Lukesch, H. (2001). *Psychologie des Lernens und Lehrens*. Regensburg: Roderer.

- Lynch, M. (2007). The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences*, 104, 8597–8604. doi:10.1073/pnas.0702207104
- Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In *Examining pedagogical content knowledge* (S. 95-132). Springer, Dordrecht. https://doi.org/10.1007/0-306-47217-1_4
- Mahler, D., Großschedl, J., & Harms, U. (2017). Using doubly latent multilevel analysis to elucidate relationships between science teachers' professional knowledge and students' performance. *International Journal of Science Education*, 39(2), 213–237. <https://doi.org/10.1080/09500693.2016.1276641>
- Maskiewicz, A. C., & Lineback, J. E. (2013). Misconceptions are “so yesterday!”. *CBE-Life Sciences Education*, 12, 352–356. doi:10.1187/cbe.13-01-0014
- Mayr, E. (1982). *The growth of biological thought: Diversity, evolution and inheritance*. Cambridge, MA: Harvard University Press.
- Moore, R., Mitchell, G., Bally, R., Inglis, M., Day, J., & Jacobs, D. (2002). Undergraduates' understanding of evolution: ascriptions of agency as a problem for student learning. *Journal of Biological Education*, 36(2), 65–71. <https://doi.org/10.1080/00219266.2002.9655803>
- Nationale Akademie der Wissenschaften Leopoldina. (2017). *Evolutionsbiologische Bildung in Schule und Universität [Teaching evolutionary biology at school and university]*. Deutsche Akademie der Naturforscher Leopoldina e.V. Nationale Akademie der Wissenschaften, Halle (Saale). Retrieved from https://www.leopoldina.org/uploads/tx_leopublication/2017_Stellungnahme_Evolutionsbiologie.pdf
- National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.
- Nehm, R. H., Beggrow, E. P., Opfer, J. E., & Ha, M. (2012). Reasoning about natural selection: diagnosing contextual competency using the ACORNS instrument. *The American Biology Teacher*, 74(2), 92–98. <https://doi.org/10.1525/abt.2012.74.2.6>
- Nehm, R. H., Poole, T. M., Lyford, M. E., Hoskins, S. G., Carruth, L., Ewers, B. E., & Colberg, P. J. (2009). Does the segregation of evolution in biology textbooks and introductory courses reinforce students' faulty mental models of biology and evolution? *Evolution: Education and Outreach*, 2(3), 527–532. <https://doi.org/10.1007/s12052-008-0100-5>

LITERATURVERZEICHNIS

- Nehm, R. H., Rector, M. A., & Ha, M. (2010). "Force-Talk" in evolutionary explanation: Metaphors and misconceptions. *Evolution: Education and Outreach*, 3, 605–613. doi: 10.1007/s12052-010-0282-5
- Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, 57(3), 263–272. <https://doi.org/10.1641/B570311>
- Nehm, R. H., & Schonfeld, I. S. (2007). Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools?. *Journal of Science Teacher Education*, 18(5), 699–723. <https://doi.org/10.1007/s10972-007-9062-7>
- Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, 45(10), 1131–1160. <https://doi.org/10.1002/tea.20251>
- Neubrand, M. (2006). Professionalität von Mathematik-Lehrerinnen und-Lehrern. *Beiträge zum Mathematikunterricht 2006*.
- Neuweg, H. G. (2004). Figuren der Relationierung von Lehrerwissen und Lehrerkönnen. In B. Hackl & H. G. Neuweg (Hrsg.), *Zur Professionalisierung pädagogischen Handelns* (S. 1–26). Münster: Waxmann.
- NGSS Lead States (2013). *Next generation science standards: for states, by states*. Washington, DC: The National Academies Press.
- Olander, C. (2012). Teaching biological evolution-internal and external evaluation of learning outcomes. *Nordic Studies in Science Education*, 5(2), 171–184. <https://doi.org/10.5617/nordina.349>
- Olszewski, J. (2010a). *The impact of physics teachers' pedagogical content knowledge on teacher action and student outcomes*. Berlin: Logos Verlag.
- Olszewski, J., Neumann, K., & Fischer, H. E. (2010b). Measuring physics teachers' declarative and procedural PCK. *Contemporary science education research: Teaching*. Ankara: PegemAkademi, 87–94.
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: knowing what students know about evolution. *Journal of Research in Science Teaching*, 49(6), 744–777. <https://doi.org/10.1002/tea.21028>
- Palmer, D. H. (1999). Exploring the link between students' scientific and nonscientific conceptions. *Science Education*, 83, 639–653. doi:10.1002/(Sici)1098-237x(199911)83:6<639::Aid-Sce1>3.0.Co;2-O

- Paris, S., Lipson, M., Wixon, K., Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8 (3), 293–316. doi: 10.1016/0361-476X(83)90018-8
- Park, S., Jang, J. Y., Chen, Y. C., & Jung, J. (2011). Is pedagogical content knowledge (PCK) necessary for reformed science teaching?: Evidence from an empirical study. *Research in Science Education*, 41(2), 245–260. <https://doi.org/10.1007/s11165-009-9163-8>
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261–284. <https://doi.org/10.1007/s11165-007-9049-6>
- Paulick, I., Großschedl, J., Harms, U., & Möller, J. (2016). Preservice teachers' professional knowledge and its relation to academic self-concept. *Journal of Teacher Education*, 67(3), 173–182. <https://doi.org/10.1177/0022487116639263>
- Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education*, 21(1), 31–61. <https://doi.org/10.1007/s12528-009-9011-x>
- Polanyi, M. (1966). *The tacit knowledge*. Garden City: Doubleday.
- Prinou, L., Halkia, L., & Skordoulis, C. (2008). What conceptions do Greek school students form about biological evolution?. *Evolution: Education and Outreach*, 1(3), 312–317. <https://doi.org/10.1007/s12052-008-0051-x>
- Reif, F. (1987). Instructional Design, cognition, and technology: applications to the teaching of science concepts. *Journal of Research in Science Teaching*, 24, 309–324. <https://doi.org/10.1002/tea.3660240405>
- Renkl, A. (2008). *Lehrbuch pädagogische Psychologie*. Bern: Huber. Verfügbar unter: <http://www.gbv.de/dms/hebis-darmstadt/toc/195706773.pdf>.
- Reynolds, A. (1992). What Is Competent Beginning Teaching? A Review of the Literature. *Review of Educational Research*, 62(1), 1–35. <https://doi.org/10.3102/00346543062001001>
- Richardson, P. W., Karabenick, S. A. & Watt, H. M. G. (Hrsg.). (2014). *Teacher motivation: Theory and practice*. New York: Routledge.
- Riese, J. & Reinhold, P. (2009). Fachbezogene Kompetenzmessung und Kompetenzentwicklung bei Lehramtsstudierenden der Physik im Vergleich verschiedener Studiengänge. *Lehrerbildung auf dem Prüfstand*, 2 (1), 104–125.

LITERATURVERZEICHNIS

- Riese, J. & Reinhold, P. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 167–187.
- Riese, J., & Reinhold, P. (2012). Die professionelle Kompetenz angehender Physiklehrkräfte in verschiedenen Ausbildungsformen. *Zeitschrift für Erziehungswissenschaft*, 15, 111–143. <https://doi.org/10.1007/s11618-012-0259-y>
- Rost, J., Prenzel, M., Carstensen, C. H., Senkbeil, M. & Groß, K. (Hrsg.). (2004). *Naturwissenschaftliche Bildung in Deutschland. Methoden und Ergebnisse von PISA 2000*. Wiesbaden: Verlag für Sozialwissenschaften.
- Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, 48(2), 117–148. <https://doi.org/10.1002/tea.20408>
- Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal*, 50(5), 1020–1049. <https://doi.org/10.3102/0002831213477680>
- Sauvé, L., Renaud, L., Kaufman, D., & Marquis, J.-S. (2007). Distinguishing between games and simulations: A systematic review. *Journal of Educational Technology and Society*, 10, 247–256.
- Schmelzing, S., Van Driel, J. H., Jüttner, M., Brandenbusch, S., Sandmann, A., & Neuhaus, B. J. (2013). Development, evaluation, and validation of a paper-and-pencil test for measuring two components of biology teachers' pedagogical content knowledge concerning the "cardiovascular system". *International Journal of Science and Mathematics Education*, 11, 1369–1390. <https://doi.org/10.1007/s10763-012-9384-6>
- Schmelzing, S., Wuesten, S., Sandmann, A., & Neuhaus, B. (2010). Measuring declarative and reflective components of biology teachers' pedagogical content knowledge. *Contemporary science education research: preservice and inservice teacher education. Ankara, Turkey: Pegem Akademi*, 71-77.
- Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology*, 46(1), 178–192. <https://doi.org/10.1037/a0016701>
- Schnotz, W. (1994). *Aufbau von Wissensstrukturen. Untersuchungen zur Kohärenzbildung bei Wissenserwerb mit Texten*. Weinheim: Beltz PVU.

- Seel, N. M. (2003). *Psychologie des Lernens*. München: Ernst Reinhardt.
- Settlage Jr, J. (1994). Conceptions of natural selection: A snapshot of the sense-making process. *Journal of Research in Science Teaching*, 31(5), 449–457. <https://doi.org/10.1002/tea.3660310503>
- Shavelson, R. J., Ruiz-Primo, M. A. & Wiley, E. W. (2005). Windows into the mind. *Higher Education*, 49(4), 413–430. doi: 10.1007/s10734-004-9448-9
- Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, 124, 209–215. doi:10.1016/j.cognition.2012.04.005
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Shulman, L. S. (1987). Knowledge and teaching. Foundations of the new reform. *Harvard Educational Review*, 57, 1–22.
- Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education and Outreach*, 1(2), 189–195. <https://doi.org/10.1007/s12052-008-0037-8>
- Southerland, S. A., Abrams, E., Cummins, C. L. & Anzelmo, J. (2001): Understanding students' explanations of biological phenomena: Conceptual frameworks or p-prims? In: *Science Education*, 85(4), 328–348. <https://doi.org/10.1002/sce.1013>
- Star, J. R., & Stylianides, G. J. (2013). Procedural and conceptual knowledge: Exploring the gap between knowledge type and knowledge quality. *Canadian Journal of Science, Mathematics and Technology Education*, 13(2), 169–181. <https://doi.org/10.1080/14926156.2013.784828>
- Steiner, G. (2001). Lernen als Wissenserwerb. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 163–202). Weinheim: PVU.
- Stover, S. K., & Mabry, M. L. (2007). Influences of Teleological and Lamarckian Thinking on Student Understanding of Natural Selection. *Bioscience*, 33(1), 11–18. <https://doi.org/10.1186/1936-6434-6-26>
- Südkamp, A. & Möller, J. (2009). Referenzgruppeneffekte im Simulierten Klassenraum. Direkte und indirekte Einschätzungen von Schülerleistungen. *Zeitschrift für Pädagogische Psychologie*, 23, 161–174. <https://doi.org/10.1037/a0027627>
- Südkamp, A., Möller, J., & Pohlmann, B. (2008). Der Simulierte Klassenraum: Eine experimentelle Untersuchung zur diagnostischen Kompetenz. *Zeitschrift für*

LITERATURVERZEICHNIS

- Pädagogische Psychologie*, 22(34), 261–276. <https://doi.org/10.1024/1010-0652.22.34.261>
- Tamir, P. (1988). Subject matter and related pedagogical knowledge in teacher education. *Teaching and teacher education*, 4(2), 99–110. [https://doi.org/10.1016/0742-051X\(88\)90011-X](https://doi.org/10.1016/0742-051X(88)90011-X)
- Tamir, P., & Zohar, A. (1991). Anthropomorphism and teleology in reasoning about biological phenomena. *Science Education*, 75(1), 57–67. <https://doi.org/10.1002/sce.3730750106>
- Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., . . . Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- Terhart, E. (1991). Pädagogisches Wissen: Überlegungen zu seiner Vielfalt, Funktion und sprachlichen Form am Beispiel des Lehrerwissens. *Zeitschrift für Pädagogik*, 27, 129–141. <https://doi.org/10.1007/s11618-017-0765-z>
- Thiel, F. (2007). Stichwort: Umgang mit Wissen. *Zeitschrift für Erziehungswissenschaft*, 10, 153–169. <https://doi.org/10.1007/s11618-007-0024-9>
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Hrsg.), *Organization of memory* (S. 382–404). New York: Academic Press.
- Van Berkum, J. J., & de Jong, T. (1991). Instructional environments for simulations. *Education and Computing*, 6, 305–358. doi:10.1016/0167-9287(91)80006-J
- Veletsianos, G., Heller, R., Overmyer, S., & Procter, M. (2010). Conversational agents in virtual worlds: Bridging disciplines. *British Journal of Educational Technology*, 41(1), 123–140. <https://10.1111/j.1467-8535.2009.01027.x>
- Voss, T. & Kunter, M. (2011). Pädagogischpsychologisches Wissen von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 193–214). Münster: Waxmann.
- Weinert, F. E. (2001a). A concept of competence: a conceptual clarification. In O. S. Rychen, L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Seattle: Hogrefe and Huber.

- Weinert, F. E. (2001b). Vergleichende Leistungsmessung in Schulen – eine umstritten Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–32). Weinheim: Beltz.
- Weitzel, H., & Gropengießer, H. (2009). Vorstellungsentwicklung zur stammesgeschichtlichen Anpassung: Wie man Lernhindernisse verstehen und förderliche Lernangebote machen kann. *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 287–305.
- Whitehead, A. N. (1929). *The aims of education*. New York: Macmillan.
- Wissenschafts- und Hochschulforschung. (2020). *Erste Förderlinie zur digitalen Hochschulbildung*. <https://www.wihoforschung.de/de/erste-foerderlinie-zur-digitalen-hochschulbildung-2378.php>
- Woolfolk Hoy, A., Davis, H. & Pape, S. J. (2006). Teacher knowledge and beliefs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of Educational Psychology* (2. Auflage, S. 715–737). Mahwah, NJ: Erlbaum.
- Yates, T. B., & Marek, E. A. (2014). Teachers teaching misconceptions: a study of factors contributing to high school biology students' acquisition of biological evolution-related misconceptions. *Evolution: Education and Outreach*, 7(1), 1–18. <https://doi.org/10.1186/s12052-014-0007-2>
- Yeh, Y.-C. (2004). Nurturing reflective teaching during critical-thinking instruction in a computer simulation program. *Computers & Education*, 42(2), 181–194. [https://doi.org/10.1016/S0360-1315\(03\)00071-X](https://doi.org/10.1016/S0360-1315(03)00071-X)
- Yip, D. Y. (1998). Identification of misconceptions in novice biology teachers and remedial strategies for improving biology learning. *International Journal of Science Education*, 20(4), 461–477. <https://doi.org/10.1080/0950069980200406>
- Zabel, J., & Gropengießer, H. (2011). Learning progress in evolution theory: climbing a ladder or roaming a landscape? *Journal of Biological Education*, 45, 143–149. doi:10.1080/00219266.2011.586714
- Ziadie, M. A., & Andrews, T. C. (2018). Moving evolution education forward: a systematic analysis of literature to identify gaps in collective knowledge for teaching. *CBE—Life Sciences Education*, 17(1). <https://doi.org/10.1187/cbe.17-08-0190>
- Zohar, A., & Ginossar, S. (1998). Lifting the taboo regarding teleology and anthropomorphism in biology education—heretical suggestions. *Science Education*, 82(6), 679–697. [https://doi.org/10.1002/\(SICI\)1098-237X\(199811\)82:6<679::AID-SCE3>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1098-237X(199811)82:6<679::AID-SCE3>3.0.CO;2-E)

ERKLÄRUNG

Hiermit erkläre ich, dass diese Dissertation – abgesehen von der Beratung durch meine Betreuerin – nach Inhalt und Form meine eigene Arbeit ist. Die Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden. Sie hat weder im Ganzen noch zum Teil an anderer Stelle im Rahmen eines Promotionsverfahrens vorgelegen. Ein akademischer Grad ist mir nicht entzogen worden. Ein Teil der Ergebnisse dieser Arbeit wurden bereits publiziert bzw. zur Publikation eingereicht (Kapitel 4 bis 6, siehe Tabelle 3.1).

Kiel, 2021

Julian Fischer

DANKSAGUNG

Meinen herzlichen Dank möchte ich an dieser Stelle Frau Prof. Dr. Ute Harms aussprechen. Sie hat mir die Möglichkeit gegeben, meine Dissertation in der Fachdidaktik Biologie anzufertigen und mich dabei jederzeit mit viel Geduld kompetent und konstruktiv beraten. Besonders möchte ich mich für das Vertrauen und den entgegengebrachten Respekt bedanken. Durch ihre kontinuierliche Förderung hat sie wesentlich zu meiner beruflichen sowie persönlichen Weiterentwicklung beigetragen.

Des Weiteren möchte ich mich bei Herrn Prof. Dr. Jens Möller bedanken, welcher mir in unseren fast wöchentlichen Treffen die Freude an pädagogisch-psychologischen Themen sowie viel über die Bedeutung und Interpretation von Zahlen vermittelte. Weiterhin möchte ich ihm für sein Interesse und die stetige Unterstützung während meines Promotionsprojektes danken.

Einen großen Dank möchte ich meinem Mentor Till Bruckermann aussprechen, der immer ein offenes Ohr für Fragen hatte und mit seinen klugen Gedanken häufig an dem Gelingen meiner Forschungsvorhaben beteiligt war. Auch die konstruktiven Gespräche und seine ausführlichen Rückmeldungen haben mich motiviert, genauer nachzudenken und einen fokussierten Blick zu behalten. Ebenso möchte ich Nils Machts danken, dass er mir bei dem Weg durch den statistischen Dschungel immer helfend zur Seite stand und mich über die komplette Promotionszeit hinweg unterstützte.

Weiterhin gilt mein Dank allen meinen Kolleginnen und Kollegen sowie den Hilfskräften, die jederzeit für ein sehr angenehmes Abreitumfeld sorgten und mich auf vielfältige Weise unterstützten.

Abschließend möchte ich mich bei meiner Frau Höpke und meinen Eltern bedanken, dass sie mir während der gesamten Promotionszeit Rückhalt gaben und niemals an mir zweifelten. Eure Liebe und Zuneigung während dieser erfahrungsreichen aber auch anstrengenden Zeit waren für mich immer ein Anker.